Implicit and explicit temporal order in the structuring of the conscious "now"

Laetitia Grabot

Université Pierre et Marie Curie

Ecole Doctorale Cerveau Cognition Comportement (ED n°158)

NeuroSpin / Cognitive Neuroimaging (Unicog)

Implicit and explicit temporal order in the structuring of the conscious “now”

Par Laetitia Grabot

Thèse de doctorat en Neurosciences Cognitives

Dirigée par Dr. Virginie van Wassenhove

Présentée et soutenue publiquement le 11/09/2017,

Devant un jury composé de :

Pr. David MELCHER	Rapporteur
Pr. Vincenzo ROMEI	Rapporteur
Pr. Lionel NACCACHE	Examinateur
Dr. Claire SERGENT	Examinatrice
Dr. Anne GIERSCH	Examinatrice
Dr. Virginie VAN WASSENHOVE	Directrice de thèse

Abstract

Time in the brain can be considered both as a dimension structuring the organization of perception and as a conscious content (van Wassenhove, 2009): this confers to time a particular flavor that makes it a paramount object of study to understand the basis of (human) cognition. The specific focus of the present thesis is to unravel the interplay between the temporal structure of perception (time as dimension) and perceived timing (time as content). From a cognitive neuroscience perspective, the purpose of my work is to bring additional empirical evidence to determine to which extent the temporal ordering of information relates to the perceived order of events, both implicitly and explicitly, and to explore the role of neural oscillations in the ordering of information. A long standing hypothesis in the field is that the brain discretizes the continuous flow of information into windows of various temporal granularities to make sense of visual or auditory information (Giraud & Poeppel, 2012; VanRullen, 2016). Hence, in the context of time perception, the notion of a "now" builds on a hierarchical organization of these temporal windows (Pöppel, 1997, 2009).

In a first line of research, we investigated how time may become an explicit conscious content by showing that temporal order is a psychological bias, which is stable over time and independent of attention. In a follow-up magnetoencephalography study, we showed that participants having a large order bias (meaning that their perceived synchrony corresponded to a large physical asynchrony) also showed the largest fluctuations in ongoing alpha power when perceived orders were contrasted for a given physical asynchrony. Alpha oscillations are herein argued to be a means to compensate for an individual's internal bias. A second line of research investigated how time implicitly shaped the perception of visual sequences. A striking consequence of discrete perception is that within the same temporal window of processing, late inputs may influence the representation of earlier information. We used the Rabbit illusion (Geldard & Sherrick, 1972) to study postdiction, in which the intermediate flash of a visual sequence is spatially mislocalized due to the temporal regularity of the sequence. We showed that when participants perceived the illusion, parieto-frontal regions were more activated following the presentation of the full visual sequence than when participants perceived the veridical visual sequence. These results suggest that high-order regions may contribute to the postdictive reconstruction of a visual sequence, consistently with the hypothesis proposing that combining uncertain sensory evidences with the expectation that stimuli tend to move slowly shapes the illusion.

Résumé

Le temps peut être à la fois une dimension structurant l'organisation de la perception et un contenu de la conscience (van Wassenhove, 2009), ce qui en fait un objet d'étude crucial pour comprendre les fondamentaux de la cognition humaine. Le but de cette thèse est d'étudier les interactions entre la structure temporelle de la perception (le temps en tant que dimension) et le temps perçu (le temps en tant que contenu). Dans une approche cognitive et neuroscientifique, le but de ce travail est de produire des preuves empiriques supplémentaires qui déterminent dans quelle mesure l'ordonnancement temporel de l'information est lié à l'ordre perçu des évènements, à la fois explicite et implicite, et d'explorer le rôle des oscillations neurales dans l'ordonnancement des évènements. Une hypothèse acceptée dans le domaine propose que le cerveau discrétise, en fenêtres de largeur temporelle variée, le flot continu de l'information visuelle et auditive pour leur donner du sens (Giraud & Poeppel, 2012; VanRullen, 2016). Dans le contexte de la perception temporelle, la notion de « présent » s'articule donc autour d'une organisation hiérarchique de ces fenêtres temporelles (Pöppel, 1997, 2009).

Le premier axe d'étude s'intéresse au temps en tant que contenu explicite et conscient et montre que l'ordre temporel est un biais psychologique, stable à travers le temps et indépendant de l'attention. Une étude complémentaire en magnétoencéphalographie a révélé que les participants ayant les plus larges biais (c.à.d. que leur perception de la simultanéité correspond à une large désynchronisation physique) montrent aussi les plus larges fluctuations de la puissance des oscillations alpha spontanées lorsque la perception de l'ordre est contrastée pour un même délai physique. Ces oscillations sont donc proposées comme moyen pour compenser un biais individuel. Le deuxième axe de recherche explore comment le temps façonne implicitement la perception d'une séquence visuelle. Une conséquence étonnante dérive de la discrétisation de la perception : au sein d'une même fenêtre d'encodage, des informations plus tardives peuvent influencer la représentation d'informations antérieures. Pour étudier ce phénomène de postdiction, nous avons utilisé l'illusion du Lapin (Geldard & Sherrick, 1972), où un flash intermédiaire, dans une séquence d'éléments lumineux, est mal localisé spatialement parlant à cause de la régularité temporelle de la séquence. Nous avons montré que percevoir l'illusion active des régions pariéto-centrales du cerveau après la fin de la séquence. Ces résultats suggèrent que des régions de haut niveau pourraient contribuer à la reconstruction postdictive d'une séquence visuelle, en accord avec l'hypothèse proposant que l'illusion dépende d'une connaissance préalable sur la vitesse d'un stimulus.

Acknowledgments

I would like to thank first Virginie van Wassenhove, my PhD advisor, for many reasons. First, thanks for having introduced me to the world of Cognitive Neuroscience, by accepting to mentor me for my M2 internship, and then for a PhD. I would also like to recognize her great management skills, always pushing when required but also encouraging so that most of the times, when leaving her office, I felt like having learned something and reinvigorated, ready to improve what needed to be. Thank you also for your support, patience and uprightness. And I will remember (among many other things): *less is more*.

Then, I would like to warmly thank all the friends and colleagues that were part of the Brain Dynamics team during my PhD for their help, the numerous discussions and laughs. Thank you to those who were here when I arrived, who help me to get acquainted with the lab life: Anne (the wrecking ball is still in the office!), Karin (thanks for all the laughs and the funny stories you always have in mind!), Nicolas, Clémence, Anja, Lucille, Luca, Milton, Ramon, Hafeza, Thiago. Thank you to those who were here almost from the beginning to the end: Baptiste (the most permanent of all the non-permanents aka senior postdoc, who can talk about things endlessly –even if some topics became recurrent-, I will proudly remember our github contribution to the lab), Benoit (who experienced with me, side by side, the ups and downs of a PhD student' life, and who will remain, above all, the best Rabbit of all time in the universe), Tadeusz (my favorite officemate despite the fact that his ideal office should be a dark and cold, windy cave), Leila (and her jokes and great mood), Ignacio (the Empanadas King), Denis (for debugging my MNE Python codes), Daria (for her kindness), Jacques (who asked me the hardest question of all time: why things never work the first time with MEG analysis?). Thank you to Sophie (for her wise advices), Renan (the French language master), Ava (the master in Bouldering organization), Pauline (I hope we gave you an engaging overview of what a cognitive neuroscientist life could be!), who arrived when I was already stuck into writing.

Many thanks to my collaborators for sharing their time, expertise and knowledge with me. I learned a lot from the methodological side from Alexandre and Tom, and shared fruitful meetings with Valérie and Lucille. Many more persons also helped me in providing a perfect work environment. I would like to thank Leila and Marco, for their availability and help with the MEG. Maryline also deserves a big thank for her efficiency in handling administrative procedures, without ever losing her smile. Thank you also to Vanna for the managing part at Unicog. Thank you to Gaelle, Laurence, and Véronique, for their invaluable help with the participants, you succeeded in making the EEG electrodes positioning part a thoroughly enjoyable moment. Thank you also to Séverine (la petite), Séverine (la grande) and Chantal for their help in MRI acquisition and their good mood. Thank you to Isabelle for her kindness and her more than welcome help with Github and Gabriel. I also received also a lot of help from the MNE-Python community, so a big thanks to all of you, especially to Alex and Denis. I would also like to thank my PhD jury for their time and their thorough questions.

I also would like to warmly acknowledge all the people from Unicog, Parietal, Unirs and Unati who were with me on this journey, and shared some conference dinners, parties, many bus trips, and a thousand of lunches at the CEA canteen: Aina, Ana-Luisa, Antonio, Andres (for your movies advices), Bianca, Catherine, Christophe, Claire, Clément (and his spinning head), Darinka, Elodie, Elvis, Esther, Evelyne, Fawzi, Florent, François (thank you for the lab retreat which left wonderful memories), Fosca, Gaël, Lucie (and your stories about crazy people), Marie (à nos aventures neurospiniennes!), Mainak, Martin (crazy discussions and saké!), Mathieu (the man who jokes faster than his shadow), Maxime (for some crazy parties), Michael, Milad (crazy tout court), Murielle, Parvaneh, Pedro, Phillippe, Sébastien, Solveig, Timo, Tomokazu, Valentina, Yoshi, Yousra, Zafer. Sorry for those whom I may have forgotten…

Last but not least, I will finish by acknowledging my close relatives, who supported me from the outside, and were already there way before the PhD, and, I hope, will still be there a long time after.

Overview

Chapter 1 provides the general framework for the empirical work addressed in my PhD thesis as well as a historical review of the notion of perceptual moments. The "now", or present moment, has been considered as a duration forming discrete time units that are hierarchically organized in the mind. One working hypothesis proposes that neural oscillations could support the temporal architecture of perception. This hypothesis has crucial implications on temporal perception. Part of my thesis work investigates whether neural oscillations may play a role in temporal order perception.

Chapter 2 investigates time as a conscious content, by focusing on explicit temporal order perception. We examine how temporal information can be translated into a conscious content. One working hypothesis is that perceived temporal order can inform on the temporal resolution of perception. Our first set of studies therefore relies on temporal order judgment tasks between simple visual and auditory stimuli. The first study behaviorally addresses temporal order perception by the prism of intra-individual variability. It directly speaks to a question asked exactly twenty years ago about the stability of time quanta (Pöppel, 1997) and reveals that temporal order is a psychological bias. The second study, using magnetoencephalography (MEG), aims at testing the hypothesis proposing that the phase of neural oscillations could encode time. We did not find any evidence in favor of a phase code for time. However, we showed that participants having a large order bias (meaning that their perceived synchrony corresponded to a large physical asynchrony) also showed the largest fluctuations in ongoing alpha power when perceived orders were contrasted for a given physical asynchrony. Alpha oscillations are herein argued to be a means to compensate for an individual's internal bias.

Chapter 3 investigates time as a structural dimension influencing perception and focuses specifically on the implicit timing governing the sequencing of visual information and the perceptual phenomenon of postdiction. We inspect how temporal information implicitly shapes conscious content. These investigations do not concern anymore the notion of time quanta but deals with higher temporal scales of processing. A striking consequence of discrete perception is that, within the same temporal window of processing, late inputs will influence the representation of earlier information. We focus on this phenomenon called postdiction in the last study combining magneto- and encephalo-graphy (M/EEG). Postdiction was advocated to explain the rabbit illusion, where a sequence of flashes is spatially misperceived due to the temporal regularity. We showed that when participants perceived the illusion, parieto-frontal regions were more activated following the presentation of the full visual sequence than when participants perceived the veridical visual sequence. These results suggest that high-order regions may contribute to the postdictive reconstruction of a visual sequence, consistently with the hypothesis proposing that combining uncertain sensory evidences with the expectation that stimuli tend to move slowly shapes the illusion.

Chapter 2 and 3 starts each with a literature review that motivates the following experimental works.

Chapter 4 mainly discusses the empirical findings reported in Chapter 2, as the work in Chapter 3 is still in progress. I discuss the origin of the individual bias in temporal order, and propose the hypothesis that attention provides a mechanism to counteract individual biases. This hypothesis has some implications on the functional role of alpha oscillations and on temporal order perception.

Part of this work has been published in international peer-reviewed journals (Chapter 2.7 and Chapter 2.9). An additional paper was submitted, as part of a collaborative work on cross-frequency coupling analysis (see Appendix). Two papers are in preparation, one related to Chapter 3.5 and one related to the collaboration on cross-frequency coupling.

Associated papers

Grabot, L., & van Wassenhove, V. (2017). 28(5):670-678. *Time order as psychological bias*. Psychological Science. doi: 10.1177/0956797616689369.

Grabot, L., Kösem A., Azizi L. & van Wassenhove V., (2017) *Prestimulus alpha oscillations and the temporal sequencing of audiovisual events.* Journal of Cognitive Neuroscience. doi: 10.1162/jocn_a_01145

Grabot, L., & van Wassenhove, V. (in prep) *Postdiction and neural oscillations.*

Appendix papers

Dupré la Tour T., Tallot L., **Grabot L.**, Doyère V., van Wassenhove V., Grenier Y., Gramfort A., *(submitted) Non-linear auto-regressive models for cross-frequency coupling in neural time series.*

Grabot, L., Kononowicz, T. W., Dupré la Tour, T., van Wassenhove, V. (in prep) *Cross-frequency coupling in time production.*

Table of Contents

Chapter 1

A historical perspective on the temporal structure of the present moment

1.1. “Now” is a retentional/protentional duration

The brain is continuously faced with an uninterrupted flow of information coming from the external world via the senses, each one sensitive to a certain kind of information. For instance, photoreceptors in the retina and hair cells in the ear are respectively transducing specific wavelength of light and specific frequency of molecule vibrations into electrical pulses. Sensing also occurs due to periodic actions such as sniffing or saccadic eye movements (Schroeder et al., 2010). Thus, sensory organs are sampling the external world, which suggests the possibility that information is discretized in the brain. Yet, the flow of consciousness appears to be perceptually continuous in account of the continuity of personal identity, as Varela (2000) was alluding to with the concept of “third level of temporality“. This apparent paradox suggests that there is a reconstruction processes at stake between the sampled information selected through the senses, and the final perception of the world. This perceptual reconstruction is not an immediate process: photoreceptors have a relaxation constant of about 35 to 75 ms (Tyler, 1985), the membrane time constant of a neuron is about 10 - 30 ms (Koch, 1996) and synaptic transmissions evolve over the second timescale (Deisz & Prince, 1989). In vision, information can take 50 to 100 ms to reach the primary visual cortex. Moreover, different features are treated in different brain areas, causing supplementary delays: color in V4/V8 areas (Hadjikhani et al, 1998), motion in V5/MT (Beckers & Zeki, 1995). The conscious *act of seeing* is definitely not immediate; it rather takes time. In his Principle of Psychology, Williams James was one of the first to propose the idea that the present is not punctual but defined by a duration, called “specious present” by E. R. Clay (James, 1890). The perceptual continuity of time constrains the specious present to be retentional and protentional (Varela, 2000 from Husserl’s terminology), meaning that the “now” contains influence of what happened in the just-past (retention), and is directed toward what is about to come (protention). To conclude about the notion of discrete time unit, here follows some words from the clear-minded review of White, written in 1963:

> *“On reviewing the ideas which have been put forth regarding the concept of a psychological unit of duration, perhaps the most striking thing to note is that although the various writers have arrived at essentially the same conclusion, they have been compelled to do so by quite different reasons. The philosophers believed that some kind of time sampling was necessary in order for one to detect change, or even to have an idea of time itself; the physiologists, as exemplified by the pacemaker theorists, believed that all vital processes are cyclic in nature, governed by basic metabolic factors; the cyberneticists, considering the brain as a computer, concluded that it must operate on a scanning principle in order to handle the volume of information it does. The fact that these different approaches have led to the same conclusion would indicate that such a conclusion deserves serious consideration.”*

1.2. Information processing is structured through hierarchical temporal windows

1.2.1. A hierarchical organization of temporal units

According to Varela (2000), perception is organized in a three-step temporal hierarchy (Fig. 1.1): the first scale corresponds to the fusion interval of each sensory system, that is to said is in the order of magnitude of the minimum necessary duration to distinguish two stimuli (2ms for audition, 100ms for vision). The second scale corresponds to a 500 ms temporal window which would reflect the "strict correlate of present-time consciousness". The last scale is at the level of narrative description. Pöppel (1972, 1997) also proposed that information processing is implemented by at least two operating levels, akin to the two first scales of Varela. The first level reflects the minimal temporal quantum, set to 30 ms, which enables the integration of signal from multiple sources by smoothing temporal contingencies. The second level is a perceptual unit binding events into a 2/3 secs sequence (Pöppel, 2009). This operating range derived from performances in duration reproduction, which decreases above this threshold, from the alternation of ambiguous percepts around this rate, and duration of working-memory (see Pöppel, 2009 for a review). Such a temporal unit would be automatically segmented and serve as a basis for conscious representation. Discretizing perception secures the maintenance and manipulation of conscious content during the present window by postponing novel percepts in the next one, but also offers enough flexibility for updating the representation with information entering in the next one. The balance between maintaining identity and enabling change of percepts is ensured by such temporal windows (Pöppel, 2009).

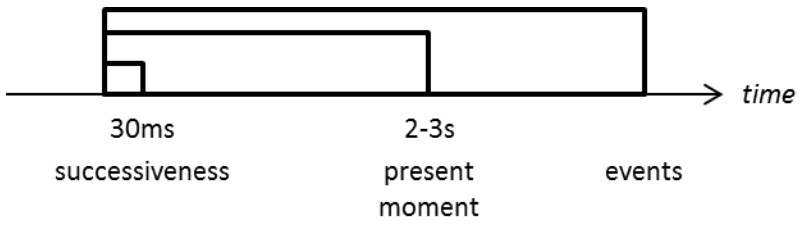


Fig. 1.1 A hierarchical and temporal organization of perception.

For larger scale (> 2-3s), the concept of perceptual events, akin to the last scale of Varela, was also proposed to represent the basic blocks of subjective experience (Zacks, 2001). Perceptual events are temporal segments with perceived beginning and ending. For instance, the different events of a typical student's day could be: "I woke up", "I ate breakfast", "I went to school". They are characterized by constancy in content: a change in environment (in direction, in motion, in time, in place, in causes…) is likely to trigger the experience of a new moment. On the contrary to Pöppel' proposal on the pre-semantic and automatic window constituting the "now", a perceptual event depends on its

representational content. This hypothesis is crucial in Kurby & Zacks' model of event segmentation (2008): they didn't search for differentiating moments, as Varela did, but put the segmentation process at the core of the mechanism, the duration of the resulting moment only being a side-effect. The event segmentation model (Kurby & Zacks, 2008) states that, as long as predictions about sensory inputs remain true (i.e. no consequent change occurs), the information is stored into working memory (WM), and as soon as the prediction error exceeds a certain threshold (i.e. a change occurs), the current representation is stored into short-term memory, which frees the WM, and a new event begins to be implemented. This means that the current event is built from all information flowing towards the brain between the moment where the previous event finishes to be implemented and the moment where the future event begins to be implemented. Because a new event is implemented when "a change occurs", this duration is not fixed, but rather depends on the environment, as well as on the demands and/or expectations of the brain regarding this environment. Moreover, events can be hierarchically organized: coarse events can be subdivided into fine-temporal grained events. Thus, event segmentation appears as a powerful way for the brain to make sense of the continuous flux of sensory inputs by simplifying and organizing, making possible to anticipate future actions. This segmentation of events seems to be a core mechanism of information processing in the brain as shown by studies measuring a decrease in reading time at events boundaries and by neuroimaging studies recording transient neural activity in temporal, parietal and frontal areas elicited at events boundaries (for review see Kurby & Zacks, 2008).

The notion of perceptual moment stems from early ideas proposed in the 19th century (see White, 1963 for a review). A non-exhaustive review of some historical psychophysics experiments is useful to understand how theories on perceptual moments were shaped and refined.

1.2.2. A historical perspective on perceptual moments

The Discrete Moment Theory (DMT, Fig. 1.2B), elaborated by Stroud in 1955, states that there is a temporal window of integration, lasting around 100 ms, within which temporal information is lost (Stroud, 1967). This hypothesis was often added to the Perceptual Intermittency Hypothesis (PIH, Fig. 1.2A), stating that the brain discretizes sensory information (Allport, 1968). One major prediction deriving from the DMT was that discrete events occurring inside a moment should be perceived as simultaneous. It was found that when letters of a word were presented within 100ms, subjects felt like the word was presented in one block, and the perception was the same even when the order of the letter was reversed (Stein & Woodworth, 1938). Lichtenstein (1961) showed that changing the inter-stimulus delay within a cycle defined by the simultaneity threshold did not influence perceived simultaneity (for review, see White, 1963).

A second prediction was that a stimulus cannot be perceived as lasting less than a minimal duration, found to be around 120 and 240 ms for visual events and 120 and 170 ms for auditory events (Efron, 1970 a and b). A third prediction of the DMT was that there is a presentation rate beyond which a train of stimuli is perceived as continuous. White (1963) was using the following task to investigate this issue: subjects had to report the number of flashes they perceived for trains of different frequencies. The

maximum perceptual rate was estimated to be at 6-7 Hz. Bartley attributed this limitation to a retinal process. White (1963) tested this hypothesis by studying electro-retinographic responses to trains of visual stimuli and results showed that the retina was able to respond to rapid sequences of stimuli up to 45 Hz (and even 125 Hz according to Riggs, 1962). A second study was undergone and elegantly discarded the proposition: trains of flashes were stereoscopically presented to the subject, which perceived a single flashing spot. The presentation rate was chosen among one leading to incorrect perception by the subject in the previous study, but so that half of this rate triggered correct perception. The results showed that subjects could indeed report accurately the correct number of stimuli in monocular viewing, and reported the same incorrect number than in the first study during stereoscopic viewing. Moreover, a similar perceptual rate was found in audition (Cheatham & White, 1954) and tactile (White & Cheatham, 1959) sensory modalities, arguing for the implication of a higher level mechanism of information integration.

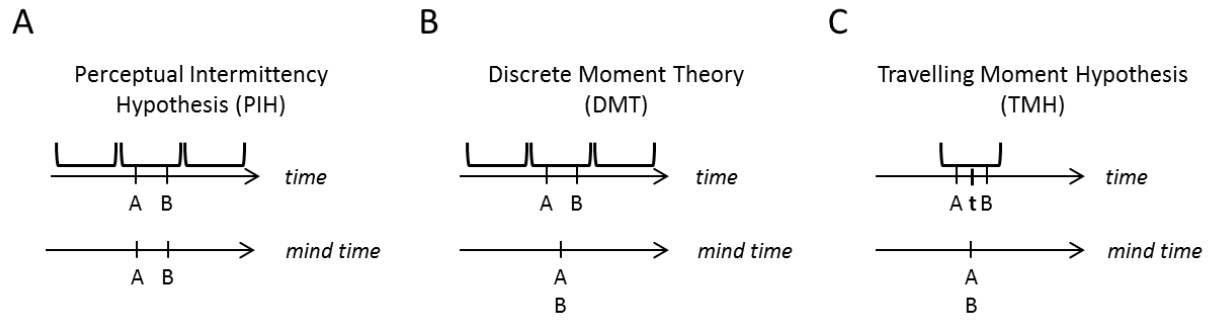


Fig. 1.2 Early theories on perceptual moments. A. The Perceptual Intermittency Hypothesis (PIH) proposed that perception is discretized. **B.** The Discrete Moment Theory (DMT) adds the assumption that temporal information is lost within a window. **C.** The Travelling Moment Hypothesis (TMH) proposed that the notion of a temporal window sliding through time.

The DMT was challenged by the Travelling Moment Hypothesis (TMH): instead of a discrete moment defined in time, the unit of processing could be sliding through time (Fig. 1.2C). At any time point, the perceptual moment is defined as the temporal segment before this time point. These two models were tested via the following task (Allport, 1968): a horizontal line occupied successively 12 positions vertically aligned on the screen (Fig. 1.3). The duration of the whole sequence, hence the rate of lines presentation, was controlled by the subject. When the duration was decreased (hence the presentation rate, increased), perception switched from a stationary array of horizontal lines with a moving dark band of shadow to a stationary array of horizontal lines flickering in synchrony. Subjects had to adjust the duration d of the sequence to reach this switch point. Two types of trials were distinguished: descending trials with large default duration, making subjects decreased it slowly, and ascending trials with small default duration, making subjects increased it slowly. In the first experiment, the sequence was presented either using the whole duration (full cycle condition), or during half of the whole duration, resulting in a twice-higher presentation rate of the 12 lines following by an empty interval (half cycle condition). The DMT predicts that if all steps in the cycle fall within one discrete moment (duration M), they will be perceived as simultaneous. For the full-cycle display, the total duration of the cycle should be d = M, for the half-cycle display, d = M (ascending trials) or d = 2M

(descending trials). On the contrary, the Travelling Moment Hypothesis predicts that all successive elements will be continuously present within the travelling moment span, i.e. d = M, in both half and full cycle conditions. Results showed that there was no significant difference in the duration between ascending and descending trials for full and half-cycle condition, speaking in favor of the Travelling Moment Hypothesis. The perceived duration was about 85 ms in full-cycle, and 95 ms in half-cycle. The second experiment gave additional evidence supporting the Travelling Moment Hypothesis: subjects had to report the direction of the moving "shadow". According to the Discrete Moment Theory, the shadow would travel in the direction opposite to that of the sequence: one position n of the sequence is off the moment, and the position n-1 will be off in the next moment. According to the Travelling Moment Hypothesis, the shadow would travel in the same direction to that of the sequence (one position n of the sequence is off the moment, and the position n+1 will be off in the next sliding moment). Results showed that subjects perceived the shadow movement in the same direction as the sequence, comforting the Travelling Moment Hypothesis.

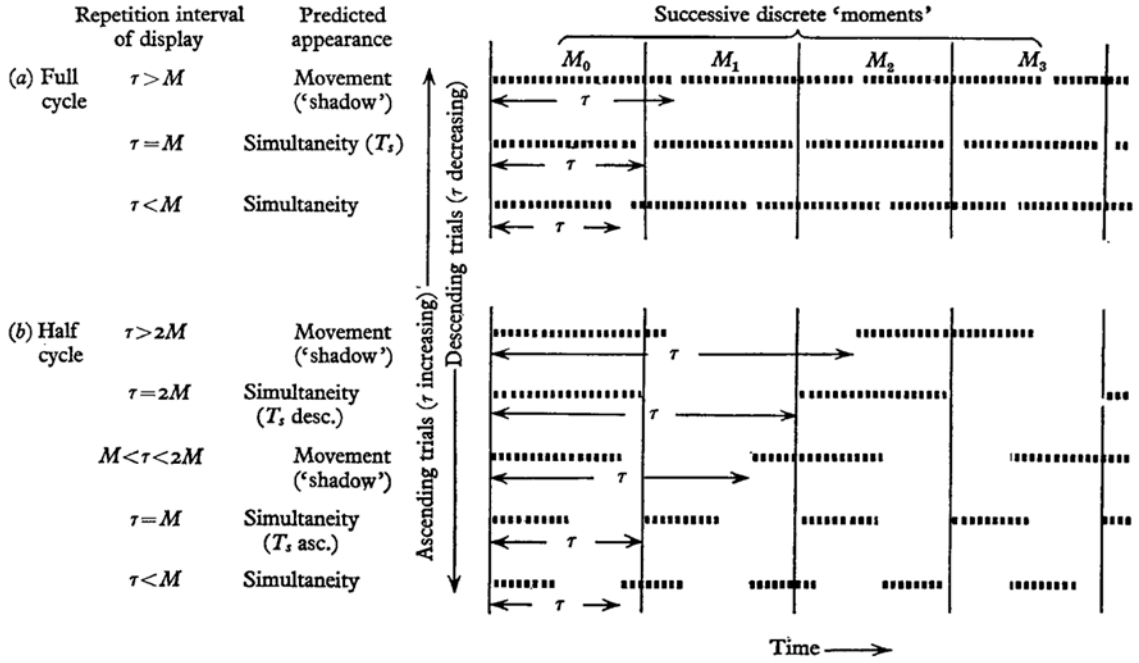


Fig. 1.3 Experimental design used to disregard the Discrete Moment Theory in favor of the Travelling Moment Theory. A horizontal line occupied successively 12 positions vertically aligned on the screen. The duration of the whole sequence (τ), so the rate of presentation of the lines, was controlled by the subject. When decreasing the duration (so increasing the presentation rate), perception switched from a stationary array of horizontal lines through which moved a dark band of shadow (called movement), to a stationary array of horizontal lines flickering in synchrony (called simultaneity). Subjects should adjust the duration of the sequence to be at this switch point (labelled M). Two types of trials were distinguished: descending trials with large default duration, making subjects decrease it slowly, and ascending trials with small default duration, making subjects increase it slowly. The sequence was presented either using the whole duration (full cycle condition), or during half of the whole duration, resulting in a twice-higher presentation rate of the 12 lines following by an empty interval (half cycle condition). The Discrete Moment Hypothesis predicts that if all steps in the cycle fall within one discrete moment, they will be perceived as simultaneous. For the full-cycle display,

the total duration of the cycle should be τ = M, for the half-cycle display, τ = M (ascending trials) or τ = 2M (descending trials). On the contrary, the Travelling Moment Hypothesis predicts that all successive elements will be continuously present within the travelling moment span, i.e. T = M in both full and half cycle conditions. (Adapted from Allport, 1968)

1.3. Neural oscillations in support of the temporal structure of perception

1.3.1. Neural oscillations as a temporal structure for perception

A neural oscillation results from the coherent activity of a neural assembly, alternating between excitatory and inhibitory cycles. Neural oscillations can be segregated into frequency bands which have been hypothesized to determine their function: delta oscillations (1-3 Hz) may internalize external rhythms to optimize perception (Lakatos et al., 2008) and mediate temporal expectations (Schroder & Lakatos, 2009; Stefanics et al., 2010); theta oscillations (4-7Hz) have been proposed to be involved in working memory processes (Lisman & Idiart, 1995); alpha oscillations (8-12 Hz) may generally reflect inhibitory processes (Klimesch et al., 2007, Jensen & Mazaheri, 2010) and may support information sampling (Varela et al., 1981; Van Rullen & Koch, 2003); beta oscillations (12-30 Hz) may underlie the maintenance of current sensorimotor goals or cognitive state (Engel & Fries, 2010) and, gamma oscillations (30-150 Hz) could reflect feature integration (Singer, 1999) and information comparison with top-down information (Herrmann et al., 2004).

The frequency division of brain rhythms, phylogenetically preserved (Buzsaki & Draguhn, 2004), indicates that the period of an oscillation could have a functional role as a temporal processing window. Larger periods allow the transfer of information across broader spatial scales while smaller periods, i.e. faster oscillations, implement local processing (von Stein & Sarnthein, 2000; Buzsaki, 2006). Synchronization of different oscillators all over the brain causes the temporal adjustment of the different processing windows which ensures an efficient communication between brain areas (Fig 1.4, Varela et al., 2001; Fries, 2005; Bastos et al., 2015). Oscillations can also be entrained by external rhythms, which provides the brain with an internalized temporal frame optimally adjusted to process environmental inputs (Lakatos et al., 2008; Schroeder & Lakatos, 2009). Last but not least, the neural oscillatory activity denotes a hierarchical organization where low frequency oscillations generally modulate higher-frequency activity (Lakatos et al., 2005). This cross-frequency coupling was empirically observed (Palva et al., 2005; Canolty et al., 2006) in crucial aspects of spatial navigation (Bragin et al., 1995; Tort et al., 2008) and working memory (Tort et al., 2009) and is thought to underlie more global mechanisms of information processing (Jensen & Colgin, 2007). The frequency hierarchy critically informs on the dual bottom-up/top-down nature of perception (Engel et al., 2001; Buschman & Miller, 2007). The brain is no longer considered as a passive device reacting to external inputs but rather as an active system self-generating information by complex loop patterns (Buzsaki, 2006, book). Perception is therefore the

product of both sensory processing and predictions, expectations, general knowledge, and contextual information carried by the ongoing activity. This bottom-up/top-down duality is reflected in the oscillatory organization, as feedback processing is conveyed by gamma oscillations while feedforward processing prevails in the alpha/beta bands (van Kerkoerle et al., 2014; Michalaeras et al., 2016).

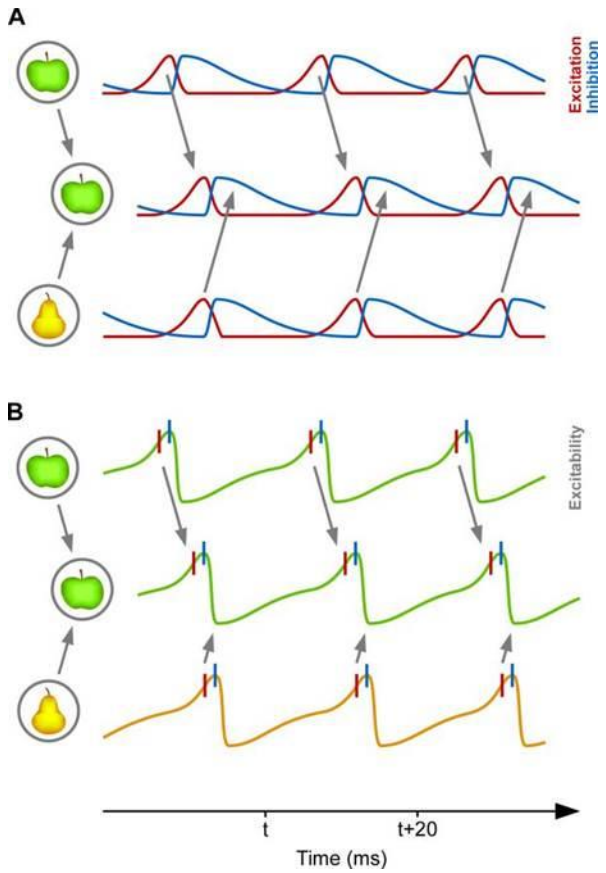


Fig. 1.4 Communication through coherence. A. The excitatory (red) and inhibitory (blue) profiles of two neuronal assemblies are depicted: on the top, one assembly codes for an apple, on the bottom, for a pear. When information is transmitted to a higher area (middle), information representing the apple arrives during an excitatory cycle in the higher area, enabling information transmission, while the pear-information is inhibited. **B.** When the excitability profile is represented, the difference between information transmission and inhibition is translated into a phase code. (Adapted from Fries, 2015)

1.3.2. A sampling window mediated by alpha oscillation

A corollary to the hierarchical structure of the present moment is the implication that perception is discretized. Cyclicality in psychophysical measurements and the disruption of perceived causality after roughly 100 ms brings empirical evidence in favor of this hypothesis (Pöppel, 1970; Dehaene, 1993; VanRullen & Koch, 2003). Given the observed time constants, the alpha oscillation has been considered a suitable candidate for implementing such a sampling window. The flickering wheel illusion offers a direct experience of the discreetness in perception: a static wheel is seen in the peripheral field as flickering at

the frequency of an individual's alpha peak (Sokoliuk & VanRullen, 2013). In visual perception, the process of individuation draws a coarse spatio-temporal object representation preceding identification. The capacity limits of 4 items (Piazza et al., 2011) does not only depend on spatial constraints but rather on temporal limitations (Wutz & Melcher, 2014). Correct individuation of items presented within 100 ms was correlated with a stronger alpha phase reset (Wutz et al., 2014), suggesting that information processing is limited to the temporal boundary of an alpha cycle.

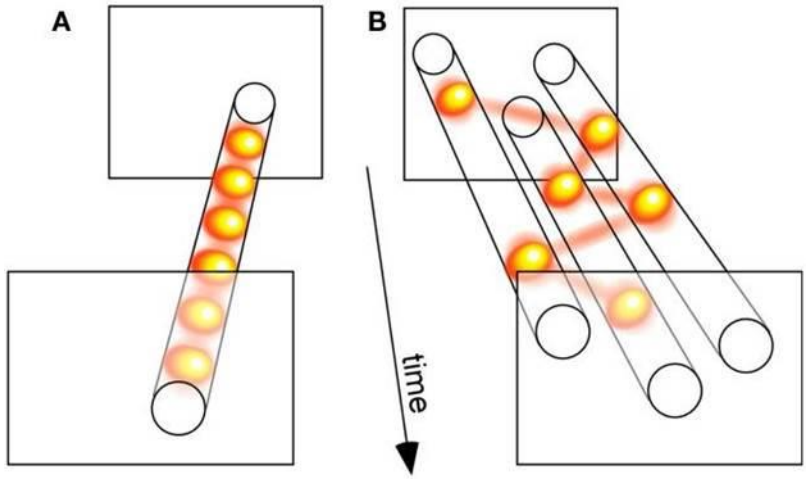


Fig. 1.5 Discrete perception and serial attention. A. Discrete perception periodicly samples information into snapshots (yellow balls). **B.** Serial attention can be seen as an evidence for discrete perception, where the attentional focus jump from an attended spatial location to another. (Adapted from VanRullen & Dubois, 2011)

On the other hand, several studies showed that the phase of ongoing alpha oscillations predicted the detectability of an event (Busch et al. 2009; Mattewson et al. 2009), by enhancing cortical excitability (Dugué et al., 2011). Thus, alpha oscillation does not only subserve perception but also anticipatory and temporal attention (Worden et al., 2000; Klimesch, 2012). From these observations, it was suggested that the mechanism discretizing perception and serial attentional processes could be two sides of the same coin (Fig. 1.5, VanRullen & Dubois, 2011). Alpha oscillations were proposed to reflect pulsed inhibition enabling the gating of sensory processing (Jensen & Mazaheri, 2010) and temporally selective activation (Klimesch, 2012). Moreover, it is very likely that the neural rhythms implementing temporal windows of processing are not unique, as foreseen by the proposal of a nested structure lying beneath the present moment (Pöppel, 1997, Varela, 2001). Rhythmic attentional sampling was also correlated with theta-modulated gamma activity (Landau et al., 2015). Modulation of perception due to the phase of pre-stimulus oscillatory activity was found for frequencies up to 30Hz, with predominance in the theta and alpha range (Fig. 1.6, VanRullen, 2016).

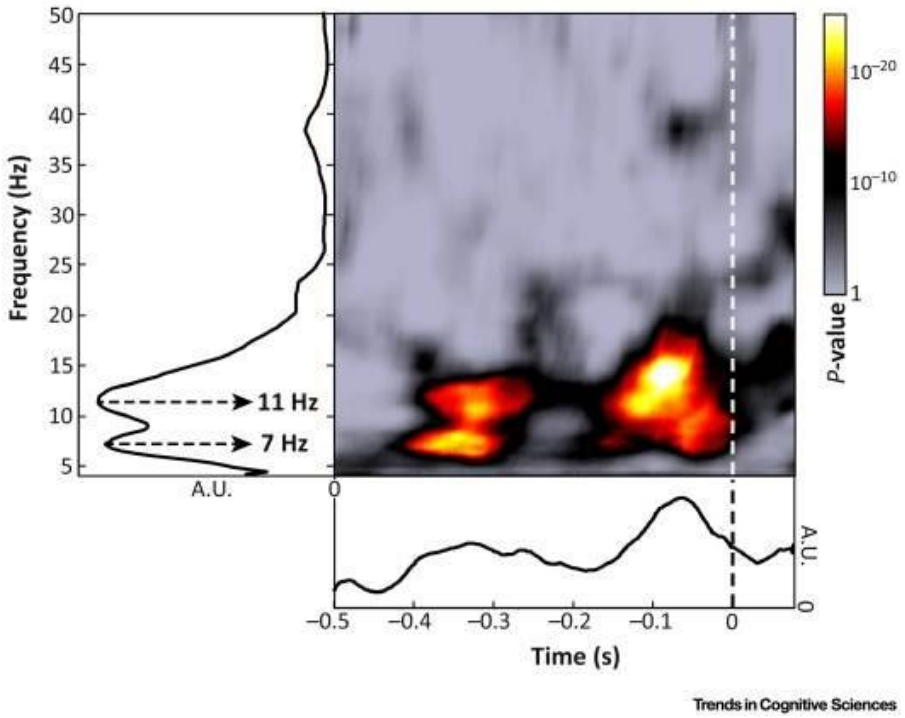


Fig. 1.6 Time-frequency measures of pre-stimulus phase divergence between two perceptual outcomes, pulled across 10 independent EEG studies. Colored values indicate significant difference of phase between the two perceptual outcomes. Two peaks in theta and alpha bands are emerging. (Adapted from VanRullen, 2016)

1.3.3. Are oscillations markers of perceptual moments?

The distinction between rhythmic and discrete perception is of fundamental importance regarding temporal perception. Rhythmic perception (Fig. 1.7A) entails that information is more efficiently processed during a specific phase of an oscillatory cycle, whereas discrete perception (Fig. 1.7B) delays processing to the next frame if a stimulus arrives too late (VanRullen, 2016). In other words, discrete perception directly guides temporal parsing, therefore influences temporal perception (Fig. 1.7D and H). One experimental prediction of discrete perception is that the oscillation frequency would affect perceived simultaneity: the wider the temporal window (i.e. the smaller the frequency), the more likely two events will fall within the same window, leading to perceiving simultaneity (Fig. 1.7 F and H). Empirical evidences indeed showed a relationship between oscillation frequency and time perception. Individual alpha frequency peak was negatively correlated with the percent of fusion percepts between two flashes (Samaha & Postle, 2015), suggesting that alpha oscillations determine the granularity of temporal perception.

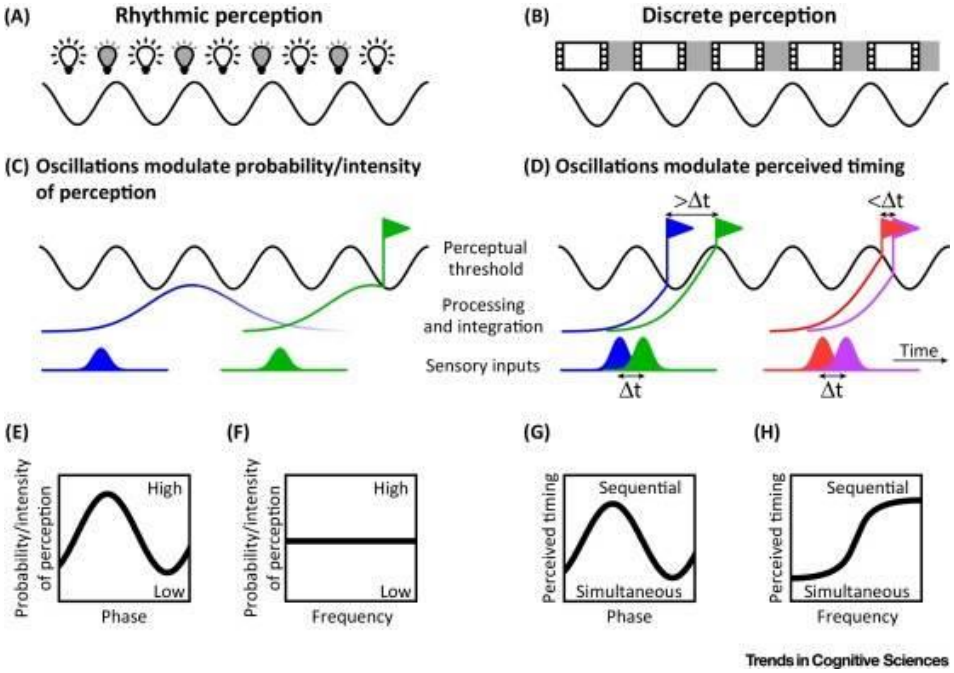


Fig. 1.7 Rhythmic versus discrete perception. A. Rhythmic perception entails that information is more efficiently processed during specific phase of a cycle (depicted by light bulbs). The oscillation (C) modulates the probability of perception (E) but a change in the frequency of the oscillation will not affect perception. **B.** Discrete perception chunks the information flow into frame or snapshot. If a stimulus arrives too late, its processing will be delayed to the next frame (D, pair of blue/green stimulus to be compared with the red/pink pair). A pair of stimulus lying in the same frame will be judged as simultaneous (red/pink pair) while a pair of stimulus processed within two frames will be perceived as sequential (blue/green pair). Perceived timing depends of the phase (G) but also of the frequency (H) of the oscillation. (Adapted from VanRullen, 2016).

A second prediction of discrete perception is that oscillation phase modulates time perception (Fig. 1.7G), which have been confirmed by empirical findings. Some studies observed that the phase of alpha oscillations predicted the perception of causality (Cravo et al., 2015) and that the phase of theta oscillations predicted the outcome of integration and segregation tasks, when trials were locked to the first stimulus onset but also to eye fixation onset (Wutz et al., 2016). This suggests that the alignment of a processing window can be triggered by external stimulation and by overt sensory sampling. Yet, while a seminal study reported that simultaneity was more likely perceived when the stimuli were presented during the occipital alpha peak than during the alpha through (Varela et al., 1981), some attempts failed to replicate these results (VanRullen & Koch, 2003). More investigations are needed to decipher whether perception is truly discrete and whether neural oscillation defined moments in temporal perception (van Wassenhove, 2017). Specifically, the question of whether the phase of neural oscillations represents conscious timing remains a burning issue (van Wassenhove, 2016).

Chapter 2

Temporal information as content in consciousness

Literature review

2.1. Perceiving temporal order

2.1.1. Temporal order as a subjective content of consciousness

Temporal order perception has intrigued scientists since the early beginning of experimental psychology, when astronomers in the 19th century discovered that a stable error when judging the timing of audition relative to vision characterized each individual – thus called the personal equation (Sanford, 1888). Temporal order is the relationship between distinct events with respect to a past/future timeline. Perceiving temporal order, a.k.a succession, signifies perceiving the sequential organization of events. Succession is one of the two concepts – with duration – that account for the notion of time (Fraisse, 1984). The twin of succession is simultaneity but both leads to uncorrelated perceptual outcomes. Indeed, the temporal order threshold i.e. the minimal duration necessary to correctly perceive the temporal order between two stimuli, is different from the simultaneity threshold, i.e. the minimal duration needed to correctly perceive simultaneity. For auditory stimuli, the temporal order threshold is about 20-40ms and the simultaneity threshold is about 5-10ms (Lackner & Teuber, 1973). Temporal order threshold appears similar across modalities for transient simple stimulation while the simultaneity threshold is larger for vision than for audition or tactile modality (Hirsh & Sherrick, 1961, Kanabus et al., 2002), suggesting that a central mechanism may be at stake in temporal order perception (van Wassenhove, 2009). Between participants, perceptual thresholds obtained in simultaneity and temporal-order judgment tasks do not correlate (Love et al., 2013; van Eijk et al., 2008), comforting the idea that the temporal representations engaged in these two tasks differs. Another alternative could be that, even if implicit timing has a fine-grained temporal resolution, the temporal resolution of conscious timing may be adjusted to task-related goals or to the complexity of information thus sometimes appears as poorer (van Wassenhove, 2009). Aiming at segregating events should equate conscious and implicit timing and consequently inform on the temporal resolution of perception.

Besides the fact that perceived order can be task- modulated, it is also subject to a large inter-individual variability. Perceiving a pair of stimuli in a different order that the one they were actually presented is a known phenomenon since the beginning of experimental psychology (Sanford, 1888) but still remains an unsolved and puzzling question. Recent evidence point to the high inter-individual variability of temporal order perception (Stone et al., 2001; Vroomen et al., 2004; Boenke et al., 2009; Love et al., 2013; Freeman et al., 2013; Kösem et al., 2014), yet, the issue of multisensory perception was rarely tackled at the subject's level although the need arises (Spence & Squire, 2003, Mattews & Meck, 2014). An interesting exception is a recent study investigating whether perceiving simultaneity is a prerequisite to achieve optimal audiovisual integration (Freeman et al., 2013). They showed that the subjective perception of simultaneity (measured by a temporal order judgment task between an

audiovisual pair of stimuli) correlates negatively with the temporal window of integration (measured by the McGurk and the stream-bounce illusion). To explain this counterintuitive finding, they proposed that the variations in timing across different mechanisms are compensated relative to their average timing, and not minimized as it was first thought. The so-called renormalization theory thus assumes that there is a quantity defined by the average time delay of temporal mechanisms in the brain that is used for compensation.

The variability in order perception between tasks and individuals points toward active mechanisms in order to achieve an optimal perception (Harris et al., 2010). Temporal order is far from being automatically encoded (Naveh-Benjamin, 1990) as neural delays and temporal noise have to be dynamically compensated for (Sugita & Suzuki, 2003).

2.1.2. Models of Temporal Order Judgment (TOJ)

To study the mechanisms underlying temporal order perception, the classical task used in the literature is the Temporal Order Judgement (TOJ) task: two near-simultaneous stimuli are presented and the observer judges which one came first. The temporal delay between the stimuli varies typically in the range of 0 to 300 ms. Analysis of responses to the range of Stimulus-Onset-Asynchronies (SOA) allows one to calculate the delay for which the responses are at chance level, i.e. the delay corresponding to the perception of simultaneity, called Point of Subjective Simultaneity (PSS). A similar task is the simultaneity judgment (SJ) task where for the same stimuli, the observer report whether he/she perceived simultaneity or not.

The first proposed model of cross-modal TOJ was the independent-channel model: each signal takes a specific time to be processed and the arrival time difference in a central mechanism is used as input of a decision function, from which is issued the final temporal order judgment (Sternberg & Knoll, 1973; Ulrich, 1987). The shape of the decision function depends on which model is considered (attention-switching, perceptual moment, triggered-moment). The PSS thus reflects the relative latency between two sensory channels, i.e. the constant error due to differences between the physical times taken by signals representing the stimuli to reach a hypothetical brain area where order is judged (Hirsh & Sherrick, 1961; Gibbon & Rutschnann, 1969; Sternberg, 1973). The so-called latency-hypothesis thus implies that other tasks estimating perceptual latencies should co-vary with the PSS, such as simultaneity judgement (SJ) and reaction time to single stimulus. However, it was largely reported that these measures were inconsistent with each other (Van Eijk et al., 2008; Love et al., 2013; Linares et al., 2014). Differences in biases were hypothesized to explain these discrepancies (Linares et al., 2014). However, whereas TOJ and RT are not consistent, modelling the latency distribution of sensory systems from reaction time to single stimulus can explain TOJ psychometric functions under the assumption that the timing component is shared and the decisional process is prone to task-specific biases (Gibbon & Rutschmann, 1969).

2.2. Temporal order and attention

2.2.1. Temporal order and attention

The temporal order judgment task was used in two different lines of research, one investigating attentional processes and one studying the temporal perception *per se*. The latter gathers studies investigating how spatial location (Zampini et al., 2003), stimulus duration (Boenke et al., 2009) or context (Vroomen et al., 2004) affect temporal order perception. For instance, when the brain is confronted to systematic asynchrony (Fig. 2.1A), recalibration causes a shift of the perceived simultaneity toward the exposure lag (Vroomen et al., 2004; Fujisaki et al., 2004; Vatakis et al., 2007; Keetels & Vroomen, 2008) and can even illusorily reverse the attribution of causality between action and perception (Stetson et al., 2006). This recalibration effect could even arise after one single audiovisual exposure (Van der Burg et al., 2013). Very few of these studies control for attention (but see Ikumi & Soto-Faraco, 2014 on the interplay between selective attention and recalibration) although perceived simultaneity is largely influenced by attention. A second important niche in the literature uses TOJ as a tool to explore the law of prior-entry (Titchener, 1908), stating that an attended stimulus is more rapidly processed and therefore will reach awareness sooner. Attention is usually manipulated toward one modality or the other (Fig. 2.1B), causing a shift in perceived simultaneity toward the non-attended modality (Sternberg & Knoll, 1973; Spence et al., 2001; Shore et al., 2001). Even if controversial evidence suggests that a shift in decisional criterion or sensory facilitation could be a better account for the shift in perceived simultaneity, than the acceleration of sensory processing, a.k.a. prior-entry (Schneider & Bavelier, 2003), carefully controlled experiments (e.g. Spence & Shore, 2001) demonstrated a true prior entry effect (Spence & Parise, 2010). Some studies also proposed that prior-entry is mediated by both a decrease in processing speed of the non-attended stimulus and a faster processing of the attended stimulus (Weiss et al., 2013, Tünnermann et al., 2015).

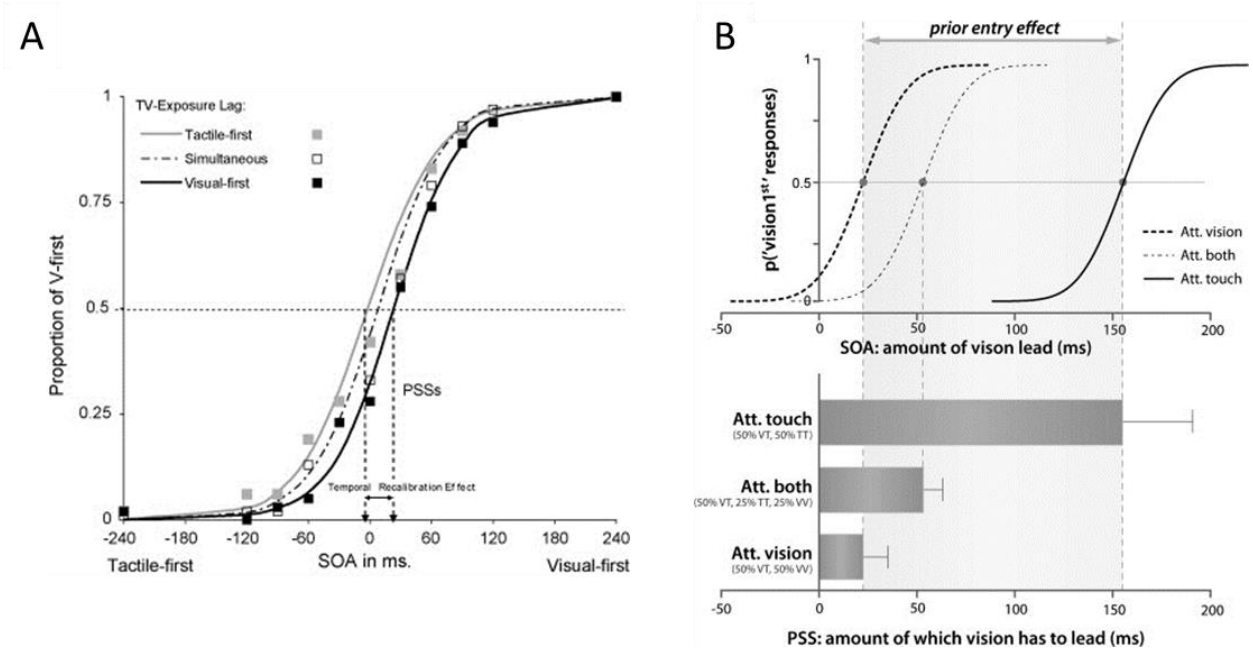


Fig. 2.1 TOJ task is used in two distinct lines of research. A. Order perception is directly studied, for instance the recalibration effect causes a shift of perceived simultaneity due to asynchrony exposure. The dashed, black and gray lines correspond to the percent of « flash first » response after respectively simultaneous, visual-first and tactile-first stimulation. **B.** A second line of work explores the law of prior-entry, stating that an attended stimulus is more quickly processed and therefore perceived earlier. The grey dashed, black and black dashed lines correspond to the percent of « flash first » response in blocks where attention was focused on both modalities, touch and vision, respectively. (Adapted from A: Keetels et al., 2008; B: Spence et al., 2010).

2.2.2. Neural correlates: spontaneous alpha fluctuations

The prior-entry states that the processing of an attended stimulus speeds up compared to a non-attended one. A direct neural corollary would be expected in the variability in latencies. Yet, no clear electrophysiological evidence were observed, as in a visuo-tactile TOJ, the prior-entry effect was associated with an earlier latency of the attended stimulus in the contralateral occipito-temporal sensors (Vibell et al., 2007) while in a visuo-tactile TOJ, it was associated with an increased amplitude in the same electrodes (MacDonald et al., 2005). Another marker of attention can be viewed in the spontaneous fluctuations of alpha power. The gating-by-inhibition theory (Fig. 2.2A) proposes that information processing is inhibited in task-irrelevant regions by means of a pulsed inhibition provided by alpha oscillations (Klimesch, 2012; Jensen & Mazaheri, 2010). Empirically, increased alpha power in irrelevant regions coupled to decreased alpha power in task-relevant regions enhances visual detection (Worden et al., 2000; Sauseng et al., 2005, Thut et al., 2006; Kelly et al., 2009; Fig. 2.2B). A decreased alpha power also predicts enhanced visual detectability (Ergenoglu et al., 2004; Hanslmayr et al., 2007) and visual discrimination (van Dijk et al., 2008). The causal relationship between alpha fluctuations, visual excitability and perception was demonstrated by, on the one hand inducing phosphenes with TMS (Romei et al., 2008) and on the other hand entraining TMS-induced occipito-parietal alpha oscillation (Romei et al., 2010). From the combination between the gating-by-inhibition framework and the law of

prior-entry, we hypothesized that pre-stimulus alpha fluctuations in sensory cortices may predict temporal order perception.

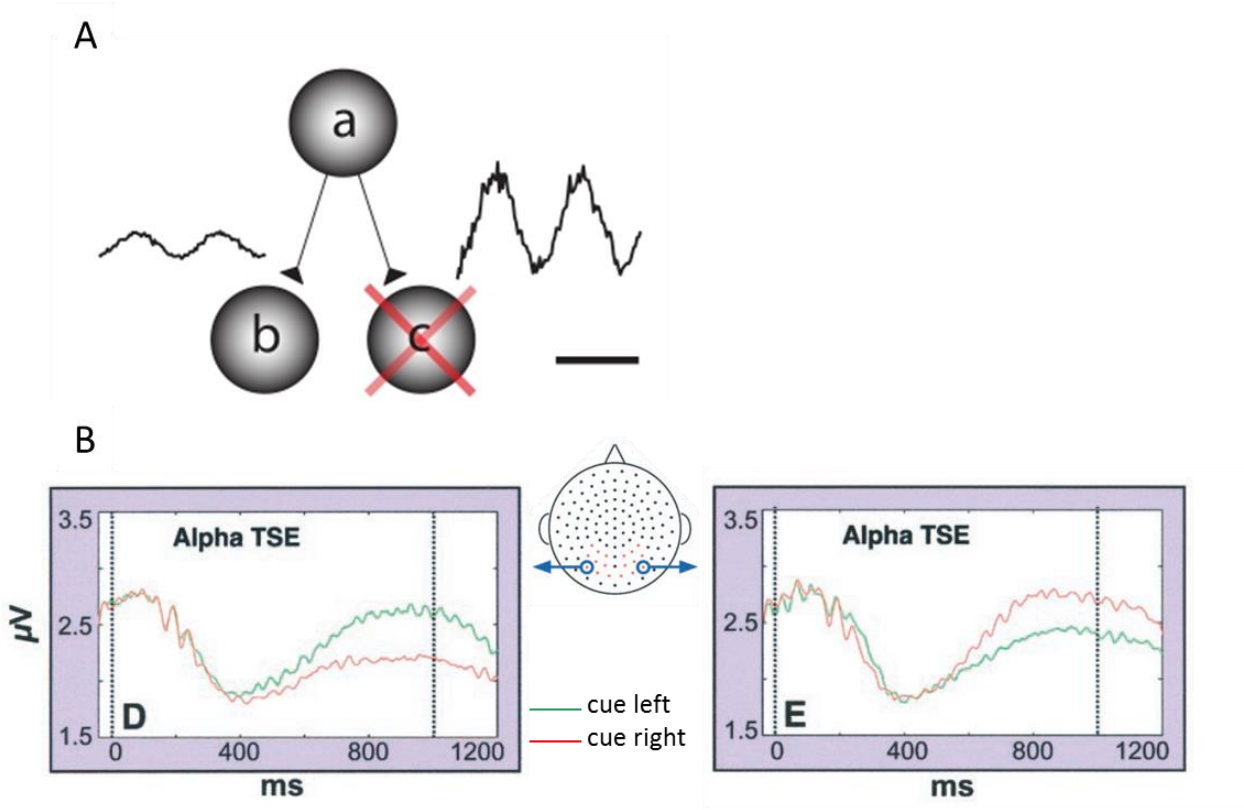


Fig. 2.2 Gating-by-inhibition A. The gating-by-inhibition framework proposes that information is gated by inhibitory pulses provided by alpha oscillation. An increase alpha oscillation will prevent information from *a* to be routed to *c*, hence favoring *a* to *b* routing. **B.** The temporal spectral evolution (TSE) of alpha oscillation is decreased in the contralateral hemisphere to the cue. The first and second vertical line represents respectively the cue and target onset. (Adapted from A: From Jensen & Mazaheri, 2010; B: Worden et al., 2000).

2.3. Neural correlates of temporal order judgment

The neural correlates of temporal order have been investigated with the direct comparison of perceived order from a TOJ task, or by comparing order perception with simultaneity perception. Other lines of work also studied order in working memory.

2.3.1. Brain regions involved in temporal order judgment

Psychophysics modelling proposed that a central system is computing the order of stimulus arrival. A potential candidate could be the parietal lobe (Battelli et al., 2007, Fig. 2.3A). TMS on the right posterior parietal cortex affected the PSS of a visual TOJ (Woo et al., 2009) and bilateral temporal

parietal junction (TPJ) was more activated in a TOJ than for a shape judgement (Davis et al, 2009; Fig. 2.3B). Correct perception of auditory order also activated bilateral TPJ but only the activity of the left TPJ was negatively correlated with individual performance (Bernasconi et al., 2010). Left post-central lesions impair auditory TOJ (Steinbuchel et al., 1999). Audiovisual asynchrony activated the right TPJ (inferior parietal lobule –IPL- and supramarginal gyrus), right DLPFC (middle frontal gyrus), left inferior parietal lobule (Adhikari et al., 2013, Fig. 2.3C) and the right insula (Bushara et al., 2001). In a tactile TOJ, the bilateral premotor cortices, middle frontal gyri, TPJ (inferior parietal cortices and supramarginal gyrus), and the superior and middle temporal gyri were more activated compared to a numerosity judgement task (Takahasi et al, 2013). Other lines of work also studied ordinality or ordering in motor sequences, which show consistent patterns with TOJ studies. Ordinality estimation was for instance associated with the activation of the middle frontal gyrus (Rubinsten et al., 2013) and serial order of motor sequence was processed in the medial premotor cortex in monkeys (Crowe et al., 2014).

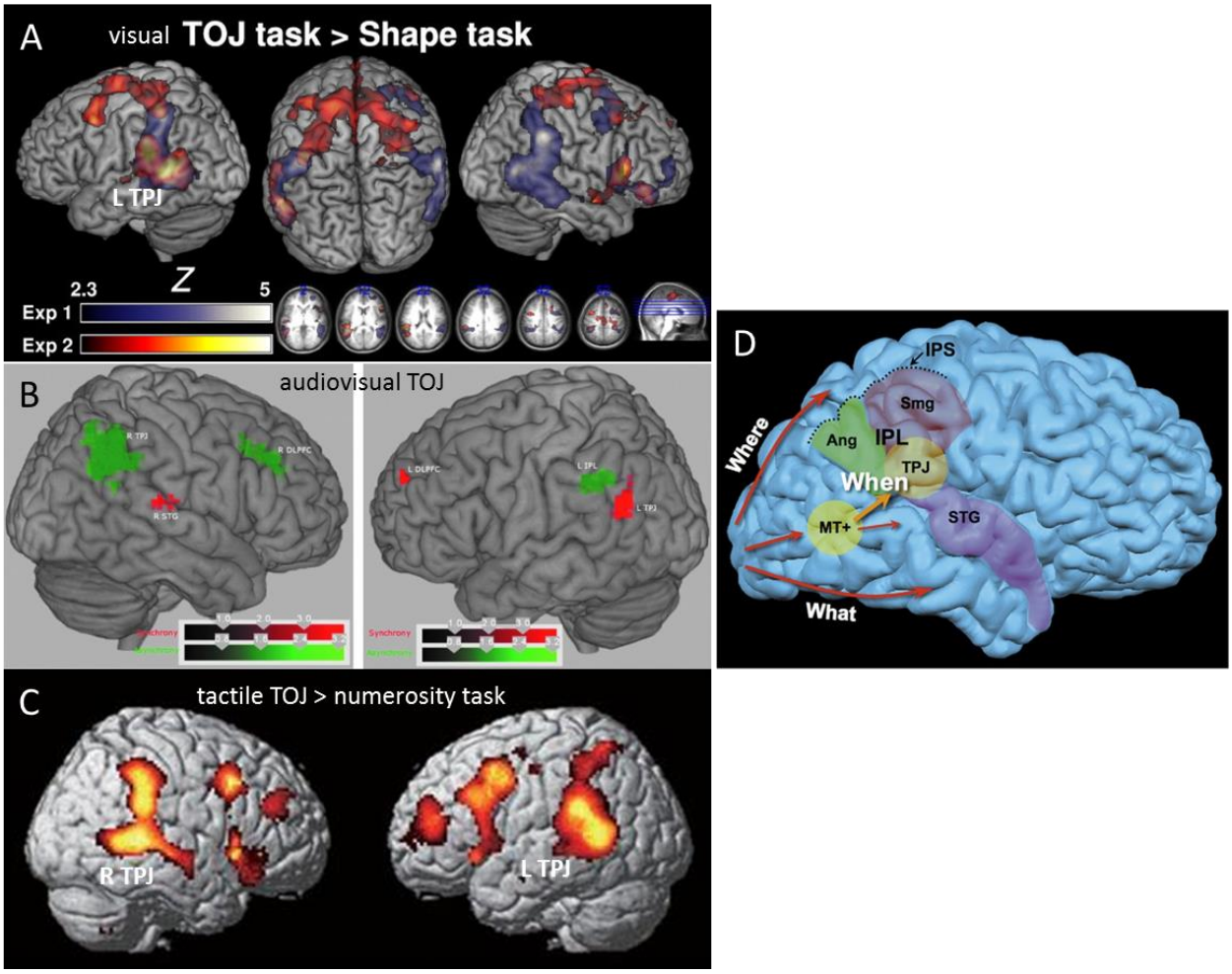


Fig. 2.3 Areas involved in visual, audiovisual and tactile TOJ. A. Bilateral TPJ and right inferior frontal gyrus and superior frontal gyrus activation was found when a visual TOJ task is contrasted to a shape discrimination task (blue, noted as experiment 1). Experiment 2 (red) controls for the potential confound that participants could attend to the stimulus onset in TOJ and not in the shape discrimination task. In experiment 2, left TPJ, the supramarginal gyrus and bilateral frontal eye fields were activated. **B.** Right TPJ

and DLPFC and left IPL (green) are significantly more activated during perception of audiovisual asynchrony compared to rest. Right STG, left DLPFC and left TPJ are significantly more activated during perception of audiovisual simultaneity compared to rest. **C.** Bilateral premotor areas, middle frontal gyrus, TPJ (inferior parietal cortices and supramarginal gyrus), were more activated in a TOJ than in a numerosity judgment task **D.** In addition to the « what » and « where » pathways respectively dedicated to object recognition and localization, a third pathway indicates « when » an object is. (Adapted from A: Davis et al., 2009; B: Adhikari et al., 2013, C: Takahashi et al., 2013; D: Battelli et al., 2007).

The neural correlates of temporal order were also studied in contrast with simultaneity perception. Audiovisual synchrony activates the superior temporal sulcus (STS, Lewis & Noppeney, 2010). Stronger connectivity from STS to primary areas indicated audiovisual coincidence (Noesselt et al., 2007). Common activations to audiovisual TOJ and SJ were found in the left inferior and superior parietal lobules, bilateral supplementary motor areas (SMA), inferior frontal gyrus and right occipital pole (Binder et al., 2015). Specific activations were found for the TOJ task only, in the left frontal-parietal networks (middle and inferior frontal gyri, superior and inferior parietal lobules) and the left occipito-temporal junction. Similar results were found in the comparison of tactile SJ and TOJ, where bilateral premotor cortex and left inferior and superior parietal lobules were more activated by TOJ (Miyazaki et al., 2016). Besides, the SMA is likely to play a central role in motor timing (Coull et al., 2011).

2.3.2. Studies of evoked-responses and pre-stimulus oscillatory activity

The dynamics of TOJ and related oscillatory activity were also investigated. Correct responses in difficult auditory TOJ were associated with smaller P2 amplitude of the evoked response (Lewandowska et al., 2008). In audiovisual TOJ, physical and perceived synchrony caused delayed auditory responses compared to physical and perceived asynchrony, while physical asynchrony triggered a delayed visual response if the percept was ambiguous compared to a correct perception (Franciotti et al., 2011). Increased amplitude of the auditory response predicts a higher accuracy in audiovisual simultaneity judgment, but only when the sound was presented first (Kaganovich & Schumaker, 2016). Altogether, these results suggest that latency differences arise asymmetrically depending on the modality.

A few studies investigated the influence of ongoing oscillations on the perception of simultaneity. Increased pre-stimulus beta power in somatosensory cortex predicts the perception of tactile simultaneity (Lange et al., 2012). Beta phase coherence in the auditory and visual cortices was also modulated by audiovisual simultaneity perception (Kambe et al., 2015). In auditory TOJ, an increased pre-stimulus beta activity in left TPJ predicts a correct order perception (Bernasconi et al., 2011). They speculated that, akin to the gating-by-inhibition framework (Jensen & Mazaheri, 2010), the beta increase functionally inhibits the neural processing of the second stimulus, decreasing the sensory load and the interference between the two sounds. It however appears an unlikely explanation for two reasons. First, beta power is rarely related to functional inhibition, but more to the maintenance of information (Engel & Fries, 2010). Then, if prestimulus beta power is indeed inhibiting information

processing, it is unclear why the processing of the second stimulus will be more impacted than the one of the first stimulus.

2.3.3. Brain regions for order in working memory

Order in working memory has been studied by two broad classes of tasks: serial recall, where a list of item is presented to a participant that have to recall them in the correct order, and order recognition on a pair of stimulus previously presented within a sequence. Behavioral evidences, such as reaction time, dissociation related to the item presentation or specific interferences, suggest that item and order recall are supported by different mechanisms and represented separately. One plausible mechanism proposed that each item is tagged by its position in the sequence, and that order is treated as a magnitude (Marshuetz, 2005). The hippocampal, parietal and prefrontal cortices were associated to order processing (for review, Marshuetz, 2005). The hippocampus was suggested to separate pattern and encode order. The order representation would then be processed in parietal regions sensitive to magnitude and prefrontal areas (dorsolateral prefrontal cortex and premotor area) would be involved in control and task-related goals. Neurons in the prefrontal cortex of monkeys indeed encoded order-based information progressively translated into goal-related representations (Genovesio et al., 2009). Additionally, the right supramarginal gyrus, the DLPFC and bilateral angular gyrus were observed during a mental time travel task (Gauthier & van Wassenhove, 2016). Altogether, these results are in line with the activation in fronto-parietal network found in perceptual temporal order studies.

2.4. The phase code hypothesis for order

Spontaneous oscillatory activity is no longer considered as a by-product of neuronal organization (Singer, 1999) but have crucial functional implications in input selection, synchronizing between-areas processing, learning (Buzsaki & Draguhn, 2004), consciousness (Engel & Singer, 2001), top-down processing (Engel et al., 2001) and information multiplexing (Akam & Kullmann, 2014). Oscillatory activity was proposed to be hierarchically organized, with the faster frequencies driven by the slower frequencies. A nested organization between delta, theta and gamma oscillations allows an optimal adjustment of the neural excitability to the timing of upcoming events (Lakatos et al., 2005). Such nested organization can be observed by cross-frequency coupling, and in particular by phase/amplitude coupling (PAC) when the phase of the slow oscillation modulates the power of the fast oscillation (Jensen & Colgin, 2007). Empirical evidences showed that theta/gamma coupling in hippocampus codes for spatial positions during rodent spatial navigation (Bragin et al., 1995). Theta phase was also proposed to index items in working memory, represented by gamma oscillations. The number of gamma bursts nested within a theta cycle could lead to the 7-item limit in working-memory capacity (Lisman & Idiart, 1995). Indeed, prefrontal theta oscillations increase during order maintenance compared to item (Hsieh et al., 2011) or spatial information (Roberts et al., 2013) maintenance. The ordering of gamma cycles predicts the perceived order of items in the parahippocampus/fusiform gyrus (Heusser et al., 2015).

During a learned sequence of pseudowords, theta/gamma coupling in auditory areas was modulated by violations of the sequence, in both monkeys and humans (Kikuchi et al., 2017).

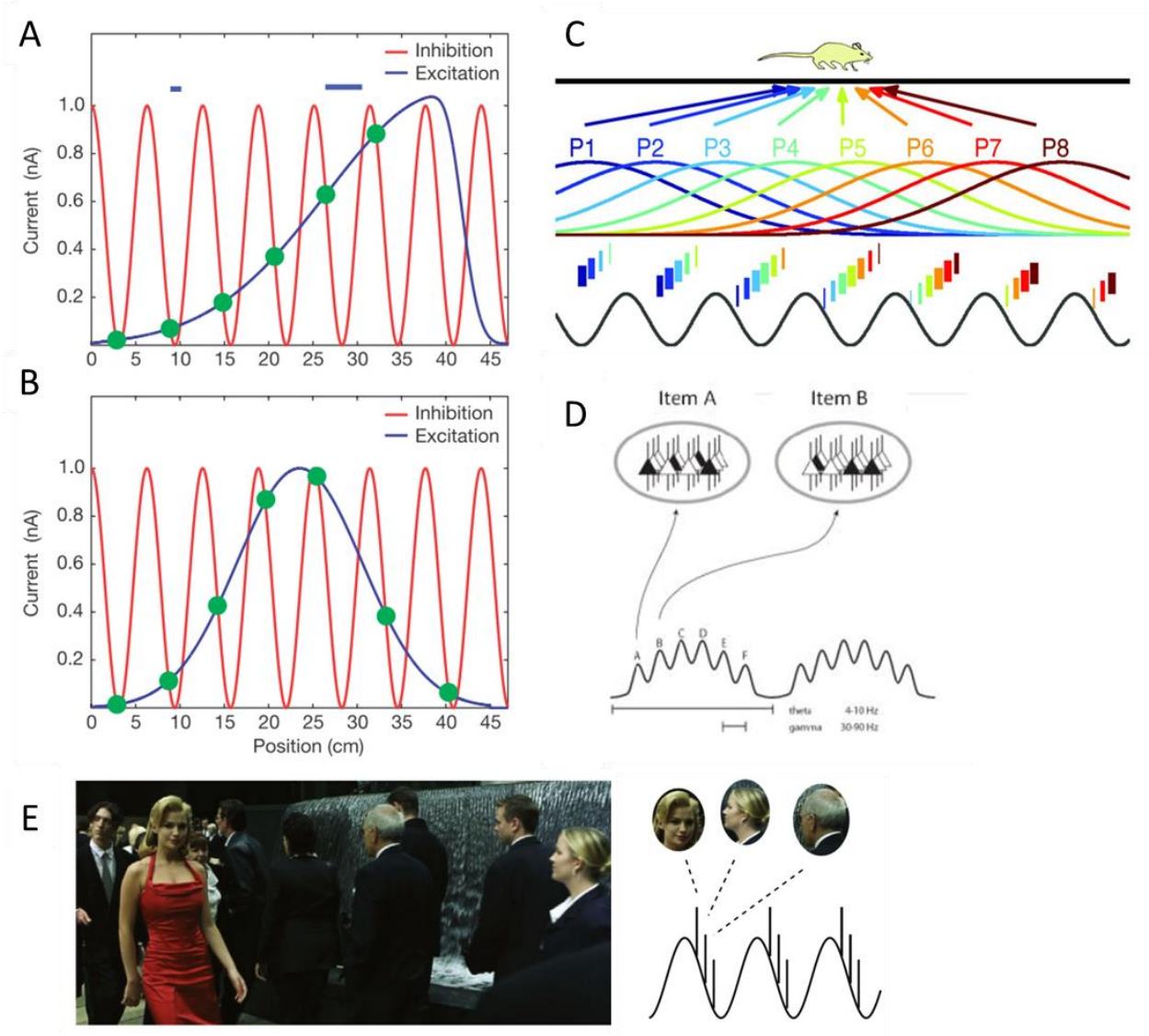


Fig. 2.4 Theta/gamma coupling, underlined by phase precession and empirically observed in spatial navigation, could form a general code to encode ordered items. A. Each neuron receives an oscillatory inhibitory input (red) and an asymmetric excitatory signal (blue). The neuron fires when excitation exceeds inhibition (green dots). When excitation increases, the excitability window is larger (blue bars) and the neurons fires earlier within the inhibitory cycles. **B.** The asymmetry of the excitatory signal is built by experience, and ensures a relationship between phase and the information represented by the firing neurons (for instance, the position of a moving rat), broken when the excitatory signal is symmetric (blue). **C.** In the rat hippocampus, each cell assembly fired at its preferential place field (P1-P8). At a specific time point, the position coded by the assembly with the strongest excitatory profile (colored lines) will be represented earlier in the theta phase (bottom dark sinusoid). The width of the colored bars indicates firing intensity of the different assemblies. **D.** The theta/gamma code could be generalized to the representation of ordered items. **E.** A speculative alpha/gamma code could prioritize the items of a visual scene in a similar way. (Adapted from A, B: Mehta & Wilson, 2002; C: Buzsaki, 2010; D: Lisman & Jensen, 2013; E: Jensen et al., 2014).

Theta/gamma coupling thus appears to support the encoding of items within a sequence, the item being a spatial position in navigation or an object in working memory. The mechanism ensuring the maintenance of the sequence order through time is called phase precession. In a coupled oscillatory system consisting of inhibitory and excitatory neurons, the increase in dendritic depolarization will cause earlier and earlier firing of spikes within the inhibitory cycles (Fig. 2.4A). In other words, the coupling strength within an excitatory network is proportional to the firing phase (Buzsaki & Draguhn, 2004). Phase precession was empirically assessed during rats' navigation (O'Keefe & Recce, 1995) where position is encoded by gamma bursts embedded within theta cycles (Fig. 2.4C). The relationship between theta phase and rats' position is stronger if the excitatory profile is asymmetric (Fig. 2.4B), which can be acquired by experience and learning (Mehta & Wilson, 2002). A sequence of events is thus compressed within a theta cycle, enabling a representation for order. The ubiquity of phase precession suggests that such oscillatory coupling could code more generally for the representation of ordered items (Lisman & Jensen, 2013, Fig. 2.4D). Another theory even proposes that such code could also represent the order in which items in a visual scene are processed (Fig. 2.4E). An alpha/gamma code was suggested to support the prioritization of items processing within a "to-do list" (Jensen et al., 2014).

These evidences suggest that the phase of slow oscillation codes for order in working memory, so we generalize this statement by hypothesizing that the phase of slow oscillation codes for perceptual order. Consistently, a recent study showed that a shift in the phase of entrained oscillations was correlated with the variability of perceived audiovisual simultaneity (Kösem et al., 2014). Additionally, the phase of ongoing alpha oscillations predicts the detectability of an event (Busch *et al.* 2009; Mattewson *et al.* 2009), by enhancing cortical excitability (Dugué et al., 2011). The phase of spontaneous alpha oscillations could thus be a possible candidate to support the representation of perceptual order.

2.5. Working hypotheses

Perceived order can be incorrect, vary between individuals and be modulated by task-related requirements on the temporal precision. This variability appears as the visible face of underlying active mechanisms overcoming neural delays and temporal noise to achieve an optimal perception of order. Temporal order judgment tasks were used to investigate temporal order perception *per se* or through the prism of prior-entry. The law of prior-entry states that an attended stimulus will reach consciousness earlier. Combined with the gating-by-inhibition framework instantiating alpha oscillations as markers of functional inhibition, we hypothesized that spontaneous alpha power in sensory cortices predicts temporal order perception. The parietal lobe including temporal parietal junction (TPJ) was commonly activated in tactile, visual, auditory and audiovisual TOJs, suggesting that it could be the central stage where order is computed. Prefrontal activity including middle frontal gyrus and/or SMA was also consistently found, potentially coordinated cognitive control. The neural oscillatory activity exhibits a hierarchical organization structuring cognitive operations. Phase-amplitude coupling enables a slow oscillation to modulate at a fine-grained temporal scale the amplitude of a faster rhythm. Theta/gamma

coupling offers a code for representing ordered items in spatial navigation but also in working memory and more speculatively in a visual scene. We hypothesized that perceptual order could be coded by the phase of slow oscillations.

An approach tackling the subjectivity of consciousness in order perception

2.6. Goal, hypothesis and summary of results

Temporal order is particularly well suited to be studied as the content of consciousness due to the high subjectivity of time perception. Yet, while subjective temporal order has been studied for decades, no clear neural correlates were found, as detailed in the previous chapter. We proposed here that some underlying neural processes could have gone unnoticed because individual differences were disregarded. Temporal order is particularly prone to substantial differences of perception between participants but, although they were reported in the literature (Sanford, 1888; Stone et al., 2001; Vroomen et al., 2004; Boenke et al., 2009; Freeman et al., 2013; Love et al., 2013; Kosëm et al., 2014;), they were not tackled as an approach to understand the neural mechanism of temporal order perception.

We hypothesized that this large inter-individual variability hides individual biases in temporal order perception. To test the stability of a potential individual bias, four sessions of an audiovisual temporal order judgement (TOJ) task were conducted within 4 months with the same participants. Results showed that each individual was indeed characterized by a stable bias in temporal order perception.

As attention is known to modulate the perception of order due to the law of prior-entry, we then questioned the relationship of this bias and attention. TOJ tasks were either used in studies investigating the effect of attention on processing facilitation (Neumann et al., 1993; Spence et al., 2001; Weiss et al., 2013), or in studies aiming to characterize other factors of temporal order perception (Boenke, 2009; Vroomen et al, 2004; Zampini et al., 2003). The former showed that the psychological measure (point of subjective simultaneity, PSS), which can be derived from TOJ tasks, was highly sensitive to manipulations of attention, while the latter were based on experimental designs that did not control for attention. We raised the question of the necessity to control for attention by asking whether the individual bias in temporal perception could be explained by attentional factors. We thus added, to the longitudinal study, conditions manipulating attention toward vision or audition. The attentional effects on PSS were replicated and findings also showed that the individual bias in temporal order perception was independent from attention. Additionally, the extent of the attentional effect was not able to fully compensate for the intrinsic bias.

These findings were published in the following article.

Grabot, L., & van Wassenhove, V. (2017). 28(5):670-678. Time order as psychological bias. Psychological Science. doi: 10.1177/0956797616689369.

2.7. Article

Research Report

Psychological Science
1–9

DOI: 10.1177/0956797616689369
www.psychologicalscience.org/PS

Time Order as Psychological Bias

Laetitia Grabot[1,2] and Virginie van Wassenhove[1,2]
[1]Cognitive Neuroimaging Unit, Institut National de la Santé et de la Recherche Médicale, Université Paris-Saclay, and [2]NeuroSpin, Commissariat à l'Energie Atomique et aux Energies Alternatives, Université Paris-Saclay

Abstract

Incorrectly perceiving the chronology of events can fundamentally alter people's understanding of the causal structure of the world. For example, when astronomers used the "eye and ear" method to locate stars, they showed systematic interindividual errors. In the current study, we showed that temporal-order perception may be considered a psychological bias that attention can modulate but not fully eradicate. According to Titchener's law of prior entry, attention prioritizes the perception of an event and thus can help compensate for possible interindividual differences in the perceived timing of an event by normalizing perception in time. In a longitudinal study, we tested the stability of participants' temporal-order perception across and within sensory modalities, together with the magnitude of the participants' prior-entry effect. All measurements showed the persistence of stable interindividual variability. Crucially, the magnitude of the prior-entry effect was insufficient to compensate for interindividual variability: Conscious time order was systematically subjective, and therefore traceable on an individual basis.

Keywords

temporal order, interindividual variability, multisensory, attention, time consciousness, open data

Received 9/17/16; Revision accepted 12/23/16

The perception of temporal order is seldom veridical (Sanford, 1888) and can be modulated by attention (Shore, Spence, & Klein, 2001; Spence, Shore, & Klein, 2001). According to Titchener's law of *prior entry* (Titchener, 1908), a stimulus to which one pays attention will be systematically perceived before a stimulus that one ignores. The law of prior entry predicts that attention prioritizes the arrival time of sensory stimuli; thus, it may govern people's subjective awareness of timing and may be instrumental in adjusting neural latencies across senses (Spence & Squire, 2003). Attention, temporal integration (Colonius & Diederich, 2004; van Wassenhove, Grant, & Poeppel, 2007), and active compensation (Sugita & Suzuki, 2003) can all contribute to even out the timing of sensory information in the brain, enabling multisensory integration. Yet even if integration yields to the compression of information in time and theoretically causes a loss of temporal resolution, the serial order of multisensory events in the few tens of milliseconds does not fully escape people's consciousness (Efron, 1973; van Eijk, Kohlrausch, Juola, & van de Par, 2008). Researchers have observed high interindividual variability in temporal-order perception (Boenke, Deliano, & Ohl, 2009; Freeman et al., 2013; Kösem, Gramfort, & van Wassenhove, 2014; Sanford, 1888; Stone et al., 2001; Vroomen, Keetels, de Gelder, & Bertelson, 2004), which suggests that attention may not be fully sufficient to compensate for possible individual biases in order perception. To date, however, no studies have tackled this question.

In the current study, we sought to determine whether participants' attentional state could fully account for interindividual variability in temporal-order perception. We used audiovisual temporal-order judgments, which allowed us to manipulate participants' attention toward one sensory

Corresponding Authors:
Laetitia Grabot, CEA, NeuroSpin—INSERM Cognitive Neuroimaging Unit, Bat 145 PC 156, F-91191, Gif-sur-Yvette, France
E-mail: laetitia.grabot@gmail.com

Virginie van Wassenhove, CEA, NeuroSpin—INSERM Cognitive Neuroimaging Unit, Bat 145 PC 156, F-91191, Gif-sur-Yvette, France
E-mail: virginie.van.wassenhove@gmail.com

modality (i.e., hearing) or another (i.e., vision). In prior work on audiovisual temporal order, the researchers have either not explicitly controlled for cross-modal attention (Boenke et al., 2009; Vroomen et al., 2004; Zampini, Shore, & Spence, 2003) or have considered the prior-entry effect to be an index of temporal-order perception (Spence et al., 2001; Weiss, Hilkenmeier, & Scharlau, 2013).

We combined both approaches in the current study and assessed audiovisual temporal-order thresholds, or audiovisual points of subjective simultaneity (PSSs), with two experimentally independent measures. The first measure of audiovisual PSS was calculated during a condition in which attention was split between hearing and vision (split-attention PSS). The second measure of audiovisual PSS was computed using two experimental conditions, one manipulating attention toward hearing and the other manipulating attention toward vision, and these conditions provided the basis for a measure free of any attentional bias (bias-free PSS). To test the stability of individual temporal-order perception over time, we used a longitudinal approach and systematically estimated both PSSs over a 4-month period. Noel, Niear, Burg, and Wallace (2016) advocated the use of longitudinal approaches to establish the robustness of multisensory timing during the life span; to the best of our knowledge, only one study (Stone et al., 2001) has reported strong within-participant correlations of audiovisual simultaneity over a 1-week period, but there was a change of viewing distance between the two sessions. The stability of PSSs and the prior-entry effect could thus be tested, along with the relationship between the prior-entry effect (i.e., the effect of paying attention to one sensory modality or the other) and PSS (i.e., a measure of temporal-order perception independent of attention to one sensory modality or the other) at the individual level. The auditory and visual PSSs and spatial biases were also assessed as a result of the spatial nature of the task that was used.

Method

Participants

Twenty-four right-handed naive participants with normal or corrected-to-normal vision and normal hearing took part in the study (9 men and 15 women; mean age = 25 years, *SD* = 4). Each provided written informed consent in accordance with the Declaration of Helsinki (World Medical Association, 2013), and the study was approved by the human research ethics committee at NeuroSpin (Gif-sur-Yvette, France). Five participants were excluded because their PSSs were located outside the range of tested stimulus onset asynchronies, or SOAs (according to the criteria used in Spence et al., 2001). Thus, in the final analysis, we included 19 participants (7 men and 12 women; mean age = 25 years, *SD* = 4), well above the requirement of 13 participants suggested by a sample-size analysis using the ICC.Sample.Size package (Rathbone, Shaw, & Kumbhare, 2015) for the R software environment (Version 3.2.4; R Development Core Team, 2016). We used the following settings: $\alpha = .05$, hypothesized value of ICC = .8, null-hypothesis value of ICC = .4, and power = .9. A post hoc power analysis revealed that our study had a power of .97 to detect the smallest ICC that we observed (.46), with $\alpha = .05$ and a null hypothesis of ICC = 0 (i.e., no intraclass reliability).

Stimuli

The experiment was designed using Psychophysics Toolbox (Version 3.0.11; http://psychtoolbox.org/) and MATLAB (Version 2014a; The MathWorks, Inc., Natick, MA). The visual stimuli were white dots (luminance = 4 lux, diameter = 1.8°) presented for 35 ms (three frames) on a black CRT screen (refresh rate = 85 Hz, resolution = 1,024 × 728 pixels). The center of the dots was presented at an eccentricity of 14° of visual angle from the horizontal meridian. Auditory stimuli were sine-wave tones (2 kHz, 63 dB) presented for 35 ms (including 5-ms fade-in and 5-ms fade-out). Speakers were placed on each side of the monitor screen at 19° of eccentricity from the center of the screen on the horizontal meridian.

The stimuli were separated by 13 different SOAs, each of which was variable: 0 ms (*SD* = 3), 34 ms (*SD* = 4), –34 ms (*SD* = 4), 54 ms (*SD* = 7), –54 ms (*SD* = 7), 68 ms (*SD* = 5), –68 ms (*SD* = 5), 98 ms (*SD* = 6), –98 ms (*SD* = 6), 145 ms (*SD* = 7), –145 ms (*SD* = 7), 237 ms (*SD* = 5), and –237 ms (*SD* = 5). A negative value for delay indicates that the left stimulus was presented first, and a positive value indicates that the right stimulus was presented first. The synchronization and stability of stimulus timing was verified with an oscilloscope using a photodiode and a microphone. Each SOA was measured 18 times for each pair of visual stimuli (flashes; VV), each pair of auditory stimuli (sounds; AA), and each pair of audiovisual stimuli (a sound and a flash; AV). A mean standard deviation of 5 ms was observed across all conditions. To ensure that the variability in stimulus delivery would not affect the measurements of individual PSSs, we modeled the distribution of possible PSSs, taking into account this variability. For this analysis, one PSS was derived from simulated data in response to a set of 13 SOAs; each of the 13 SOAs was assigned the highest possible standard deviation measured in the stimulus delivery (i.e., *SD* = 7 ms). This procedure was repeated a thousand times, which allowed us to have a surrogate distribution of simulated PSSs based on the observed variability in our stimulus delivery. The mean difference between the surrogate distribution and the veridical PSS computed with no-delay SOAs was 0.2

ms (*SD* = 2.4), 95% confidence interval, or CI = [−3.7, 4.1]. All PSS estimations reported in the Results section are thus robust, and estimation errors are largely below the observed and reported individual differences in the study.

Procedure

The longitudinal study was composed of four sessions: the first session and three follow-up sessions that took place 7 days (*SD* = 2), 31 days (*SD* = 3), and 124 days (*SD* = 8) after the first session. A participant's perception of temporal order was assessed using a lateralized temporal-order-judgment task (Shore et al., 2001). The task took place in a darkened room, and each participant sat with his or her head on a chin-rest located 65 cm away from a computer screen. On any given trial, the participant was presented with a pair of stimuli: AA, VV, or AV; (Fig. 1a). The two stimuli were segregated laterally—one on the left side, one on the right side—independently of their sensory modality. Participants were asked to favor accuracy over speed. They reported whether they perceived the left or right stimulus first in a two-alternative forced choice. Before the experiment, participants were trained with 14 trials at maximal SOAs.

In addition, participants' attention to sensory modalities was manipulated in a manner orthogonal to the requirements of the temporal-order-judgment task (Spence & Parise, 2010) using three attention conditions (Fig. 1b). In auditory-attention blocks, participants were asked to pay attention to the sounds only; in visual-attention blocks, to pay attention to flashes only; and in split-attention blocks, to split attention between sounds and flashes. At the beginning of each block, a letter centered on the screen informed participants which sensory modality they should attend to ("A" for auditory, "V" for visual, and "AV" for split attention). In the split-attention blocks, visual and auditory stimuli were equally likely, so that six AV trials (three with the sound on the left, three with the sound on the right), three AA trials, and three VV trials were presented for each SOA. In the auditory-attention blocks, sounds were presented in 66% of the trials so that eight AV and four AA trials were presented for each SOA; in the visual-attention blocks, flashes were presented in 66% of the trials so that eight AV and four VV trials were presented for each SOA (see Table S1 in the Supplemental Material available online). No AA trials were presented in the visual-attention blocks and no VV trials were presented in the auditory-attention blocks. There were a total of 10 blocks per experimental session (three auditory-attention blocks, three visual-attention blocks, and four split-attention blocks), making a total of 1,560 trials per session. The blocks were presented in random order and separated by rest periods. One experimental session lasted about 75 min in total.

PSS and ICC

For each of the four experimental sessions, the percentage of right-stimulus-first responses was plotted as a function of the SOA in the auditory-attention and visual-attention conditions; in the split-attention condition, right-stimulus-first responses were first converted to flash-first responses. To establish the individual's psychometric curve per condition and per session, data were fitted to binomial distributions using a probit link function estimated with a generalized linear model in MATLAB. Goodness of fit (R^2) was .90 on average and above .73 in 95% of the fits. The temporal-order threshold, or PSS, was defined as the SOA for which a participant's performance was at chance level—that is, at 50% of flash-first responses for the PSS estimating audiovisual temporal order (Fig. 1c, top) and 50% of right-stimulus-first responses for the PSSs estimating auditory temporal order and visual temporal order.

Three different audiovisual PSSs could be estimated in this experiment (Fig. 1b): split-attention PSS, auditory-attention PSS, and visual-attention PSS. To determine whether the split-attention PSS was subject to an individual's attentional bias toward one or the other sensory modality, we computed a bias-free PSS measure as the average of the audiovisual PSSs in the visual- and the auditory-attention conditions.

The rationale for the bias-free PSS was that biases toward one sensory modality should be maximal in the visual- and auditory-attention conditions and should thus cancel out when averaged. Therefore, the bias-free PSS was considered a measure of audiovisual PSS that was free of any attentional biases. The existence of an individual bias independent of top-down attentional strategy in temporal-order estimation should lead to a stable bias-free PSS. If the split-attention PSS and bias-free PSS significantly and systematically differed, this would further substantiate that there is a need to control participants' attentional focus in future studies of audiovisual temporal-order perception.

To assess the stability and the reliability of an individual's PSS across sessions, we used the Irr package (Garner, Lemon, Fellows, & Singh, 2012) for the R software environment. We calculated the ICC, which is a statistical measure of whether intra- or interclass variability best accounts for the variance in the data. In this case, the class was the individual, so the intraclass variability was the measure of variance of PSS across sessions within an individual, and the interclass variability was a measure of variance across individuals and sessions. ICCs were calculated as the ratio of the interindividual variability to the sum of the intraindividual, interindividual, and noise variabilities. A one-way model quantification of the consistency of measures based on single ratings was used (Hallgren, 2012).

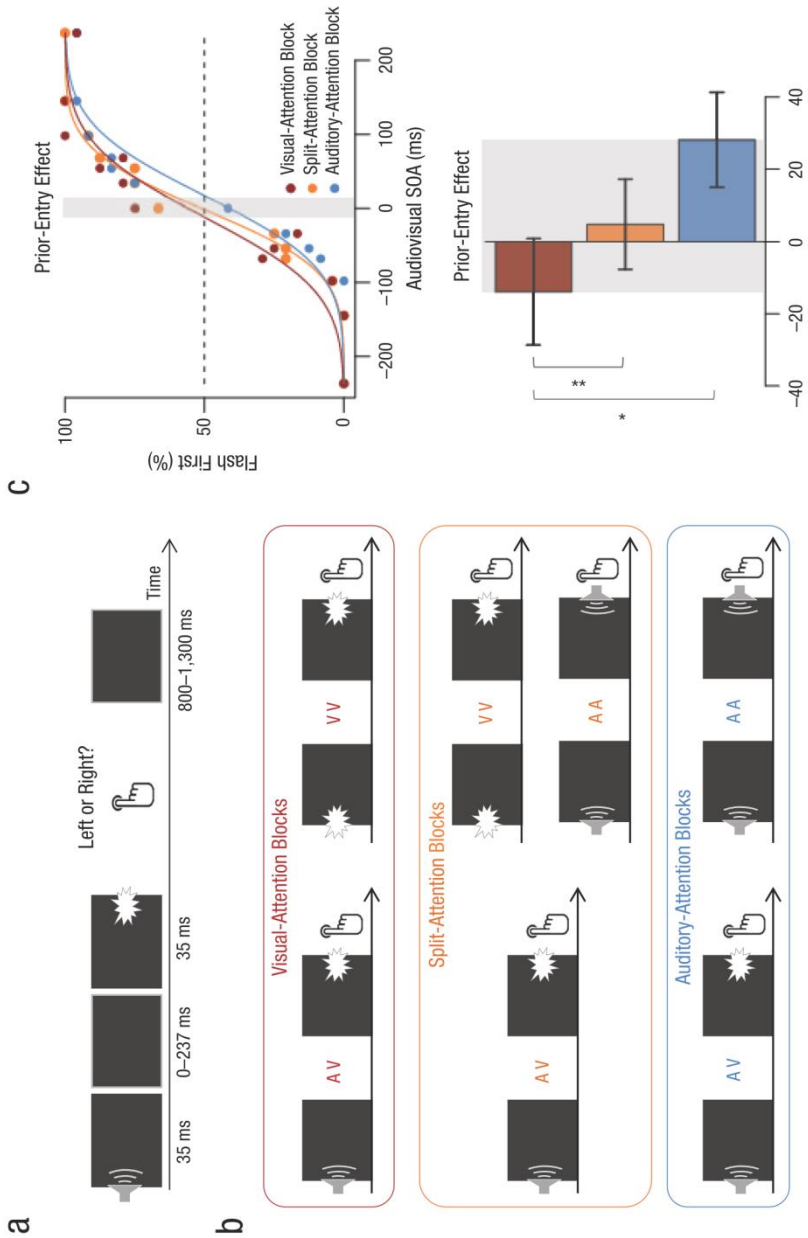


Fig. 1. Experimental design and results. In each trial of the spatial temporal-order-judgment task (a), participants were presented with a pair of lateralized stimuli (e.g., a sound on the left and a flash on the right, as illustrated here). Participants reported whether they perceived the first stimulus to have occurred on the left or on the right, regardless of its sensory modality. Performance in three experimental conditions (b) was tested to control for attentional biases in participants' point of subjective simultaneity (PSS): In visual-attention blocks, participants attended to visual events (flashes) during audiovisual (AV) or visual-only (VV) trials. In split-attention blocks, participants paid equal attention to auditory events (sounds) and visual events during AV, VV, or auditory-only (AA) trials. In auditory-attention blocks, participants attended to auditory events during AV or AA trials. Responses in the split-attention condition were sorted according to which sensory modality was perceived first, which yielded an estimate of the probability of reporting having seen a flash first at each stimulus onset asynchrony (SOA). The top graph in (c) presents the percentage of trials in each attention condition on which a representative participant saw a flash first as a function of the audiovisual SOA. The bottom graph in (c) presents grand-mean PSSs, estimated on an individual basis and averaged across participants separately for each attention condition. The gray shaded areas in the graphs indicate the magnitude of the prior-entry effect. Error bars indicate ±1 *SEM*. Asterisks indicate significant differences between conditions ($^{*}p < .05$, $^{**}p < .01$).

Results

Prior-entry effect

Based on the previous literature (Spence et al., 2001), our expectation was that the audiovisual PSSs observed during the different attention conditions would shift toward the nonattended sensory modality. For instance, during auditory attention, we expected that a flash would have to be presented earlier to be perceived as having occurred simultaneously with the sound, yielding a more positive PSS; conversely, during visual attention, we expected that a sound would have to be presented earlier to be perceived as having occurred simultaneously with the flash, yielding a more negative PSS. The audiovisual PSSs obtained in each attention condition were submitted to two-way analyses of variance (ANOVAs) with factors of attention (auditory attention, visual attention, split attention), session (first session, 1 week later, 2 weeks later, 4 months later) and nested random effects modeling the between-participants variability.

As predicted by the law of prior entry, we found a main effect of attention, $F(2, 18) = 6.12$, $p = .006$, $\eta_p^2 = .290$, which supports the observation that the PSS observed in split attention shifted toward positive values during auditory attention and negative values during visual attention (Fig. 1c, bottom half). A post hoc Bonferroni-corrected paired t test indicated that the audiovisual PSS during visual attention was significantly smaller ($M = -14$ ms, $SD = 3$ ms) than that during auditory attention ($M = 28$ ms, $SD = 4$), $t(15) = -2.72$, $p = .047$, $d = 0.75$, and that during split attention ($M = 5$ ms, $SD = 3$), $t(15) = -4.05$, $p = .003$, $d = 0.34$. However, the audiovisual PSSs during auditory attention and split attention did not differ statistically, $t(15) = -1.77$, $p > .250$, $d = 0.45$. We found no significant effect of session, $F(3, 18) = 0.70$, $p > .250$, $\eta_p^2 = .044$, and no significant interaction between attention and session, $F(6, 18) = 1.27$, $p > .250$, $\eta_p^2 = .078$, which suggests that the modulations of PSS under various attention conditions remained stable across sessions and therefore over time. Although participants were asked to favor accuracy over speed, we performed an analysis of reaction times (RT) as a function of attention and sensory modalities: Results showed that the prior-entry law did not predict the differences of RT in temporal-order judgment and that RTs in the split-attention condition were overall faster when the sound was presented first (see Fig. S1a in the Supplemental Material). These analyses also suggested that RTs were largely decorrelated from PSS, as reported in previous studies (Jaśkowski, 1999).

Stability of the PSS across sessions

The stability of the PSS across experimental sessions was assessed using ICC. The ICC for split-attention PSS was .69, 95% CI = [.50, .85], $F(18, 57) = 10.10$, $p < .001$, which means that 69% of the observed variance was due to interparticipant variability and 31% was due to intraparticipant variability. This result suggests that the split-attention PSSs were reliable across the four experimental sessions (Fig. 2a). Likewise, bias-free PSS remained stable over time, ICC = .77, 95% CI = [.61, .89], $F(18, 57) = 14.70$, $p < .001$ (Fig. 2b). We thus questioned the link between the two estimates of audiovisual PSS, considering that sorting the individual values for split-attention PSS and bias-free PSS showed very similar ranking (Fig. 2c). The correlation between split-attention PSS and bias-free PSS was highly significant, $r = .92$, 95% CI = [.81, .97], $t(17) = 9.99$, $p < .001$. The coefficients of the linear regression of the split-attention PSS on the bias-free PSS were 1.02 for the slope and −2.10 for the intercept, which indicates that the two PSS measures were approximately the same (Fig. 2c, inset). These results strongly suggest that split-attention PSS was a reliable measure of individual bias in a temporal-order-judgment task and that it seemed immune to participants' attention strategy. The measures of the auditory-attention PSS and the visual-attention PSS showed stability over time as well, along with participants' bias to the left side in these tasks (see Fig. S2 in the Supplemental Material). It is noteworthy that participants' introspective reports on task difficulty rated the AV temporal-order judgment to be more difficult than the AA and VV temporal-order judgments, and the attention conditions were rated as being equally difficult (see Fig. S3 in the Supplemental Material).

Linking the prior-entry effect and PSS

To further investigate the relationship between PSS and the prior-entry effect, we quantified the magnitude of the prior-entry effect by subtracting PSS during visual attention from PSS during auditory attention and tracked its stability over time. The findings suggest that the magnitude of the prior-entry effect was also stable within individuals over the 4-month period, across-session ICC = .46, 95% CI = [.23, .70], $F(18, 57) = 4.41$, $p < .001$ (Fig. 3a). Hence, there was a large interindividual variability in the prior-entry effect, but it remained stable over time. In particular, 3 (out of 19) participants showed negative prior-entry effects in at least two of the four sessions; among the remaining participants, six additional prior-entry values were negative. This suggests either that nonattended events were prioritized or that participants did not correctly follow the task instructions. The inclusion or exclusion of these data did not affect the main results of our previous or subsequent analyses, so they were retained.

The extent of the prior-entry effect can be seen as the range of possible values an individual's PSS took as a function of the individual's attentional focus (Fig. 3b). Considering that one functional role of prior entry could be compensation for temporal delays to achieve veridical simultaneity (or simultaneity constancy; e.g., Kopinska &

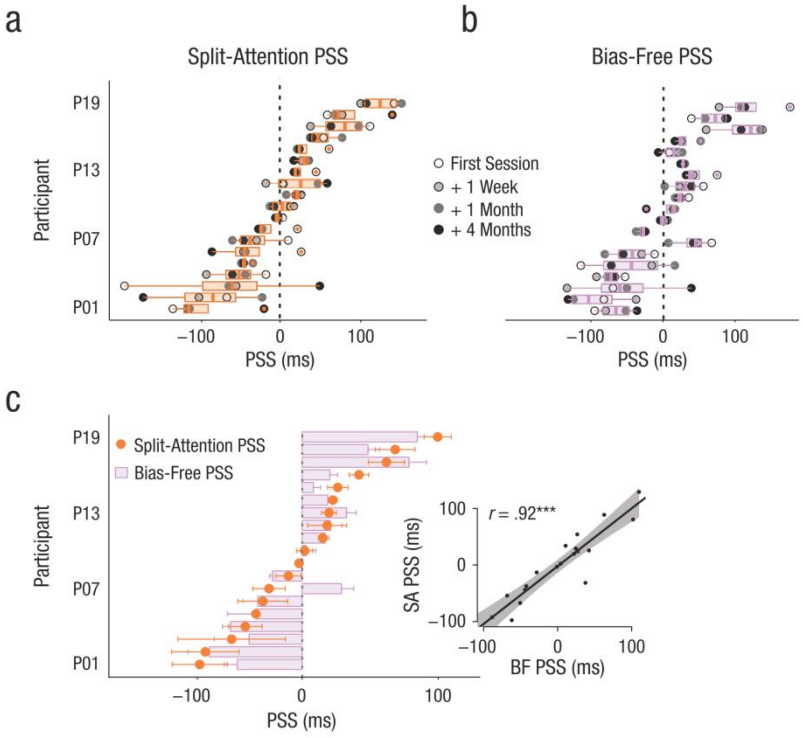


Fig. 2. Points of subjective simultaneity (PSSs) for individual participants. The graphs in (a) and (b) show results for the (a) split-attention condition (split-attention PSS) and (b) visual- and auditory-attention conditions combined (bias-free PSS). Each row is for an individual participant. The circles indicate results for the individual sessions, and the box-and-whiskers plots show the within-participant variance across all four sessions. The left and right edges of the boxes represent the first and third quartiles, respectively, and the lines down the center of the boxes represent the medians. The left and right ends of the whiskers represent the minimum and maximum PSSs, respectively. Participants were sorted by the PSS required for the participant to perceive audio and visual stimuli simultaneously, from most negative (i.e., the sound had to be presented first) at the bottom to most positive (i.e., the visual event had to be presented first) at the top. The circles with a different color in the center represent outliers. In (c), both the split-attention PSS and the bias-free PSS are presented for each participant. The results are ordered from the smallest (bottom) to the largest (top) split-attention PSS value. The orange error bars indicate ±1 *SEM*; the purple error bars indicate +1 *SEM*. In some cases, the error bars are too small to be seen here. The scatterplot (with best-fitting regression line) in the inset shows the relationship between the split-attention PSS (SA PSS) and the bias-free PSS (BF PSS). The shaded area indicates the 95% confidence interval. The asterisks indicate a significant correlation between the two types of PSS (***$p < .001$).

Harris, 2004), individuals with a PSS far from veridical simultaneity were predicted to display larger magnitudes of the prior-entry effect. To test this prediction, we calculated the correlation between the magnitude of the prior-entry effect (Fig. 3b, shaded area) and the absolute value of the split-attention PSS values (Fig. 2c) for each individual and for each session. The two were significantly correlated, $r = .34$, 95% CI = [.13, .53], $t(74) = 3.16$, $p = .002$, which suggests that participants with a larger PSS magnitude tended to exhibit larger prior-entry effects (Fig. 3c).

Discussion

In a longitudinal study, we showed the existence of a robust and stable interindividual variability in audiovisual temporal-order perception (and similar variability in auditory and visual temporal-order perception). Our results suggest that temporal order is a psychological bias that is unique to each individual and may result from structural constraints (such as the intrinsic neural delays of a given individual's brain; Freeman et al., 2013) or

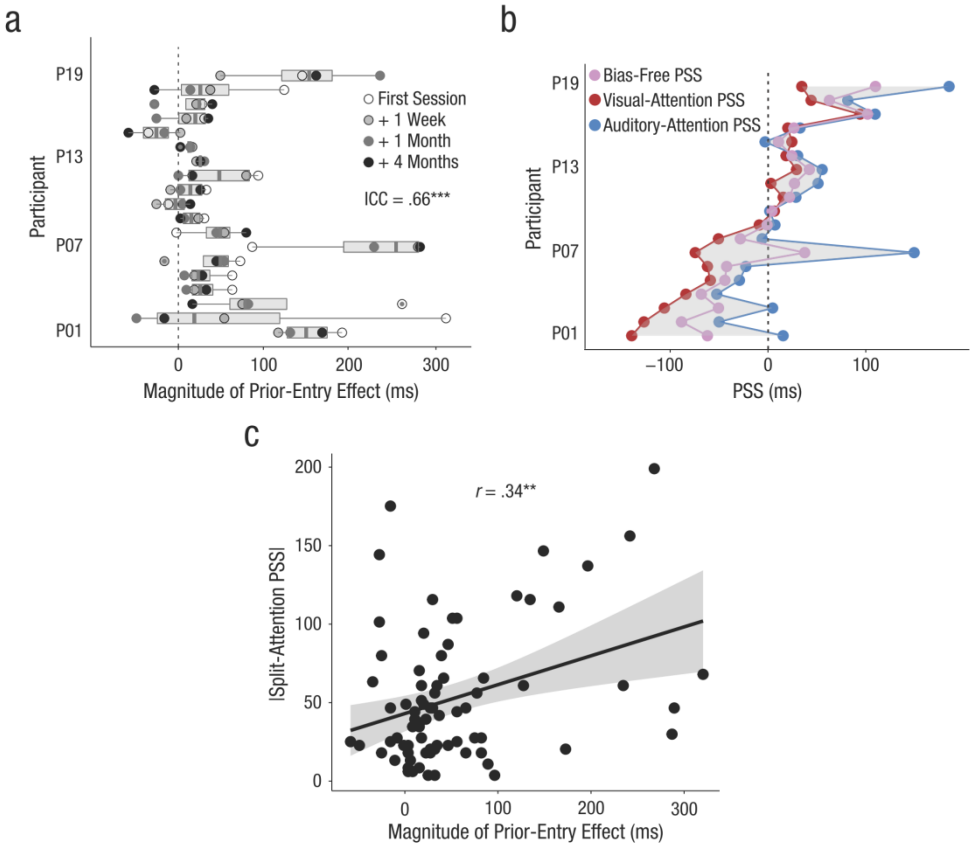


Fig. 3. The prior-entry effect for individual participants. The graph in (a) shows the prior-entry effect observed for audiovisual stimuli, computed from visual- and auditory-attention conditions. The circles indicate results for the individual sessions, and the box-and-whiskers plots show the within-participant variance across all four sessions. The left and right edges of the boxes represent the first and third quartiles, respectively, and the lines down the center of the boxes represent the medians. The left and right ends of the whiskers represent the minimum and maximum scores, respectively. The asterisks indicate a significant intraclass correlation coefficient (***p < .001). The graph in (b) shows each participant's average bias-free, visual-attention, and auditory-attention point of subjective simultaneity (PSS) over the four experimental sessions. The shaded area represents the magnitude of the prior-entry effect. The scatterplot (with best-fitting regression line) in (c) shows the relationship between the absolute value of the PSS in the split-attention condition and the magnitude of the prior-entry effect. Each dot corresponds to an individual at a single session. The shaded area represents the 95% confidence interval. The asterisks indicate a significant correlation (**p < .01).

from functional constraints akin to hidden-state variables (Wexler, Duyck, & Mamassian, 2015).

Interindividual variability in audiovisual temporal-order perception was observed regardless of whether participants split attention across senses (split-attention PSS) or not (bias-free PSS): The two indices were significantly correlated, which indicates that individual temporal-order bias persisted through manipulation of attention. When participants attended to one or the other sensory modality, shifts in the audiovisual PSS predictably followed the law of prior entry but fluctuated around the individual bias-free PSS. These shifts were consistent over time and opposite in direction to those of the attended sensory modality. We also found that the magnitude of the prior-entry effect was significantly correlated with the absolute value of the split-attention PSS: the further away the individual bias from physical simultaneity, the larger the magnitude of the prior-entry effect. One possible

functional interpretation of this finding is that prior entry may reflect brain mechanisms that compensate for individual temporal-order bias or Bayesian priors. In other words, our results suggest that attention mediates the prior-entry effect, enabling an individual's brain to compensate to some extent for its intrinsic temporal biases.

The existence of a psychological bias in temporal order is consistent with observations that PSSs measured for various stimulus classes (beeps and flashes, audiovisual speech) are correlated within individuals (Love, Petrini, Cheng, & Pollick, 2013) despite differences in integration time (Vatakis & Spence, 2010). Thus, intrinsic biases may depend not on low-level stimulus features but on high-order brain computations, consistent with how temporal-order judgments typically activate parietal cortices (Adhikari, Goshorn, Lamichhane, & Dhamala, 2013; Battelli, Pascual-Leone, & Cavanagh, 2007; Davis, Christie, & Rorden, 2009; Woo, Kim, & Lee, 2009). The extent to which such temporal-order bias plays a role in the integration of information, however, remains unclear: If an individual's subjective simultaneity has been shown to correlate with the size of the temporal-integration windows in audiovisual speech (Freeman et al., 2013), the PSSs obtained in simultaneity and temporal-order judgment tasks do not correlate (Love et al., 2013; van Eijk et al., 2008). These two measures probably capture different processes, and temporal-integration windows computed across individuals may be confounded by interindividual variability in temporal order. Taken together, these results emphasize the importance of quantifying interindividual differences in psychological and neuroimaging studies (Kanai & Rees, 2011) to shed light on possibly unnoticed neural mechanisms underlying temporal cognition.

Action Editor

Philippe G. Schyns served as action editor for this article.

Author Contributions

L. Grabot and V. van Wassenhove designed the study. L. Grabot scripted the experiment, collected the data, and performed the data analysis and interpretation under the supervision of V. van Wassenhove. L. Grabot and V. van Wassenhove wrote the manuscript. Both authors approved the final version of the manuscript for submission.

Acknowledgments

We thank members of NeuroImagerie Applicative Clinique et Translationnelle for their help in recruiting participants. Preliminary results were presented at the annual meeting of the Association for Psychological Science, in Chicago, May 2016.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was supported by European Research Council Starting Grant ERC-YStG-263584 (to V. van Wassenhove).

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797616689369

Open Practices

All data have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/6n894/. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797616689369. This article has received the badge for Open Data. More information about the Open Practices badges can be found at https://www.psychologicalscience.org/publications/badges.

References

Adhikari, B. M., Goshorn, E. S., Lamichhane, B., & Dhamala, M. (2013). Temporal-order judgment of audiovisual events involves network activity between parietal and prefrontal cortices. *Brain Connectivity*, *3*, 536–545. doi:10.1089/brain.2013.0163

Battelli, L., Pascual-Leone, A., & Cavanagh, P. (2007). The '*when*' pathway of the right parietal lobe. *Trends in Cognitive Sciences*, *11*, 204–210. doi:10.1016/j.tics.2007.03.001

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

Colonius, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: A time-window-of-integration model. *Journal of Cognitive Neuroscience*, *16*, 1000–1009.

Davis, B., Christie, J., & Rorden, C. (2009). Temporal order judgments activate temporal parietal junction. *Journal of Neuroscience*, *29*, 3182–3188. doi:10.1523/JNEUROSCI.5793-08.2009

Efron, R. (1973). Conservation of temporal information by perceptual systems. *Perception & Psychophysics*, *14*, 518–530.

Freeman, E. D., Ipse, A., Palmbaha, A., Paunoiu, D., Brown, P., Lambert, C., . . . Driver, J. (2013). Sight and sound out of synch: Fragmentation and renormalisation of audiovisual integration and subjective timing. *Cortex*, *49*, 2875–2887. doi:10.1016/j.cortex.2013.03.006

Garner, M., Lemon, J., Fellows, I., & Singh, P. (2012). Irr: Various coefficients of interrater reliability and agreement (Version 0.84) [Software]. Retrieved from https://cran.r-project.org/package=irr

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorial in Quantitative Methods for Psychology*, *8*(1), 23–34. doi:10.20982/tqmp.08.1.p023

Jaśkowski, P. (1999). Reaction time and temporal-order judgment as measures of perceptual latency: The problem of dissociations. In G. Aschersleben, T. Bachmann, & J.

Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 265–282). Amsterdam, The Netherlands: North-Holland/Elsevier Science.

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*, *12*, 232–242. doi:10.1038/nrn3000

Kopinska, A., & Harris, L. R. (2004). Simultaneity constancy. *Perception*, *33*, 1049–1060. doi:10.1068/p5169

Kösem, A., Gramfort, A., & van Wassenhove, V. (2014). Encoding of event timing in the phase of neural oscillations. *NeuroImage*, *92*, 274–284. doi:10.1016/j.neuroimage.2014.02.010

Love, S. A., Petrini, K., Cheng, A., & Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PLoS ONE*, *8*(1), Article e0054798. doi:10.1371/journal.pone.0054798

Noel, J.-P., Niear, M. D., Burg, E. V., & Wallace, M. T. (2016). Audiovisual simultaneity judgment and rapid recalibration throughout the lifespan. *PLoS ONE*, *11*(8), Article e0161698. doi:10.1371/journal.pone.0161698

Rathbone, A., Shaw, S., & Kumbhare, D. (2015). ICC.Sample.Size: Calculation of sample size and power for ICC (Version 1.0) [Software]. Retrieved from http://cran.r-project.org/package=ICC.Sample.Size

R Development Core Team. (2016). R: A language and environment for statistical computing (Version 3.2.4) [Computer software]. Retrieved from https://www.r-project.org/index.html

Sanford, E. C. (1888). Personal equation. *American Journal of Psychology*, *2*, 3–38.

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychological Science*, *12*, 205–212. doi:10.1111/1467-9280.00337

Spence, C., & Parise, C. (2010). Prior-entry: A review. *Consciousness and Cognition*, *19*, 364–379. doi:10.1016/j.concog.2009.12.001

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*, 799–832.

Spence, C., & Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Current Biology*, *13*, R519–R521. doi:10.1016/S0960-9822(03)00445-7

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., . . . Porter, N. R. (2001). When is now? Perception of simultaneity. *Proceedings of the Royal Society B: Biological Sciences*, *268*, 31–38. doi:10.1098/rspb.2000.1326

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*, *421*, 911. doi:10.1038/421911a

Titchener, E. B. (1908). *Lectures on the elementary psychology of feeling and attention*. New York, NY: Macmillan.

van Eijk, R. L., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception & Psychophysics*, *70*, 955–968. doi:10.3758/PP.70.6.955

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45*, 598–607. doi:10.1016/j.neuropsychologia.2006.01.001

Vatakis, A., & Spence, C. (2010). Audiovisual temporal integration for complex speech, object-action, animal call, and musical stimuli. In M. J. Naumer & J. Kaiser (Eds.), *Multisensory object perception in the primate brain* (pp. 95–121). New York, NY: Springer.

Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research*, *22*, 32–35. doi:10.1016/j.cogbrainres.2004.07.003

Weiss, K., Hilkenmeier, F., & Scharlau, I. (2013). Attention and the speed of information processing: Posterior entry for unattended stimuli instead of prior entry for attended stimuli. *PLoS ONE*, *8*(1), Article e0054257. doi:10.1371/journal.pone.0054257

Wexler, M., Duyck, M., & Mamassian, P. (2015). Persistent states in vision break universality and time invariance. *Proceedings of the National Academy of Sciences, USA*, *112*, 14990–14995. doi:10.1073/pnas.1508847112

Woo, S.-H., Kim, K.-H., & Lee, K.-M. (2009). The role of the right posterior parietal cortex in temporal order judgment. *Brain and Cognition*, *69*, 337–343. doi:10.1016/j.bandc.2008.08.006

World Medical Association. (2013). *WMA Declaration of Helsinki: Ethical principles for medical research involving human subjects*. Retrieved from http://www.wma.net/en/30publications/10policies/b3/

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental Brain Research*, *152*, 198–210. doi:10.1007/s00221-003-1536-z

Neural data supports psychological findings

2.8. Goal, hypothesis and summary of results

The phase of neural oscillations is a suitable candidate to encode time as the content of consciousness, as argued in the introductory section. Phase indeed reflects perceived order when oscillations are entrained by an external rhythm (Kösem et al., 2014). We aimed to test whether oscillatory phase can explain temporal order perception in the absence of entrainment. As a ubiquitous rhythm in the brain, the alpha oscillation appears as a possible candidate to encode temporal order (see the previous section for justifications). In an MEG study, we tested whether the alpha phase just before the presentation of an audiovisual pair of stimuli could predict the perceived order, in absence of any entrainment. We did not find any evidence in that direction. We thus proposed that the phase of the oscillation may reflect timing – the temporal structure of events - but not time – as the content of consciousness.

An alternative hypothesis was built on the gating-by-inhibition framework (Jensen & Mazaheri, 2010). High alpha power was related to the inhibition of information processing, which increased performance when distractor information was efficiently inhibited. We hypothesized that high auditory alpha power coupled to low visual alpha power would predict the perception of the visual stimulus first, whereas a high visual alpha power coupled to a low auditory alpha power would predict the perception of the auditory stimulus first. While classical group analysis between perceived orders did not reveal any significant differences in pre-stimulus alpha power, at the individual level, decreased alpha power predicted a correct response, whose modality depended on the participant's PSS. For a negative-PSS participant, a correct response was to perceive the sound first, while for a positive-PSS participant, a correct response was the perception of the flash first. Crucially, the alpha power difference between perceived orders was linearly correlated with PSS. A larger PSS was associated with a larger alpha power difference between the perceptions of the flash versus the sound first, whereas individual with a PSS near to physical simultaneity showed almost no alpha power differences between perceptual judgments.

This findings stress the importance of to taking inter-individual differences into account, which can reveal mechanisms that would have otherwise been overlooked. They are highly consistent with the existence of a psychological bias in the temporal order perception system. The alpha oscillation appears to gate the weights of sensory evidence and of an intrinsic bias, tipping the scales in favor of a correct or biased perceptual decision.

These findings were published in the following article.

Grabot L, Kösem A., Azizi L. & van Wassenhove V., (2017) *Prestimulus alpha oscillations and the temporal sequencing of audiovisual events.* Journal of Cognitive Neuroscience. doi: 10.1162/jocn_a_01145

2.9. Article

Prestimulus Alpha Oscillations and the Temporal Sequencing of Audio-visual Events

Laetitia Grabot[1], Anne Kösem[2,3], Leila Azizi[1], and Virginie van Wassenhove[1]

Abstract

■ Perceiving the temporal order of sensory events typically depends on participants' attentional state, thus likely on the endogenous fluctuations of brain activity. Using magnetoencephalography, we sought to determine whether spontaneous brain oscillations could disambiguate the perceived order of auditory and visual events presented in close temporal proximity, that is, at the individual's perceptual order threshold (Point of Subjective Simultaneity [PSS]). Two neural responses were found to index an individual's temporal order perception when contrasting brain activity as a function of perceived order (i.e., perceiving the sound first vs. perceiving the visual event first) given the same physical audiovisual sequence. First, average differences in prestimulus auditory alpha power indicated perceiving the correct ordering of audiovisual events irrespective of which sensory modality came first: a relatively low alpha power indicated perceiving auditory or visual first as a function of the actual sequence order. Additionally, the relative changes in the amplitude of the auditory (but not visual) evoked responses were correlated with participant's correct performance. Crucially, the sign of the magnitude difference in prestimulus alpha power and evoked responses between perceived audiovisual orders correlated with an individual's PSS. Taken together, our results suggest that spontaneous oscillatory activity cannot disambiguate subjective temporal order without prior knowledge of the individual's bias toward perceiving one or the other sensory modality first. Altogether, our results suggest that, under high perceptual uncertainty, the magnitude of prestimulus alpha (de)synchronization indicates the amount of compensation needed to overcome an individual's prior in the serial ordering and temporal sequencing of information. ■

INTRODUCTION

To build a stable representation of the world, the integration of information within and across sensory modalities is essential. Multisensory information is typically integrated within a few tens of milliseconds (Spence & Parise, 2010; Vroomen & Keetels, 2010; van Wassenhove, Grant, & Poeppel, 2007), consistent with the range of temporal tuning observed in multisensory neuronal populations (Meredith, Nemitz, & Stein, 1987; Benevento, Fallom, Davis, & Rezak, 1977). Within such timescales, the temporal sequencing of sensory information is a priori irrelevant to establish a stable representation of a conscious integrated content (Pöppel, 2009; Pöppel, Schill, & von Steinbüchel, 1990); indeed, temporal resolution would be lost by virtue of integration. In this context, it is surprising that temporal order perception—that is, the conscious representation of primitive temporal sequences—is highly variable across individuals (Grabot & van Wassenhove, 2017; Ipser et al., 2017; Kösem, Gramfort, & van Wassenhove, 2014; Hanson, Heron, & Whitaker, 2008; Fujisaki, Shimojo, Kashino, & Nishida, 2004; Vroomen, Keetels, de Gelder, & Bertelson, 2004; Stone et al., 2001), possibly indicating individual specificity in the duration of information processing (Stone et al., 2001). Consistent with this, temporal order perception displays large interindividual variability but within-individual stability over weeks (Grabot & van Wassenhove, 2017). If the outcome of multisensory integration informs on the content of conscious representations, the perceived timing of sensory information may thus inform on the structuration of conscious representations (Lashley, 1951).

In this magnetoencephalography (MEG) study, we investigated the neural sources of interindividual variability in audiovisual temporal order perception and specifically asked whether ongoing intrinsic neural fluctuations could predict the perception of audiovisual temporal order. Participants were presented with near threshold audiovisual stimuli (individually tailored) while they performed a temporal order judgment (TOJ) task. The temporal asynchrony between the auditory and visual stimuli was chosen on a per individual basis so as to yield ambiguous order perception and to contrast brain responses to two differently perceived sequences given the same physical stimulation—that is, audio perceived first, thereafter referred to as "AV," or visual perceived first, thereafter "VA."

[1]Université Paris-Sud, Université Paris-Saclay, Gif/Yvette, France, [2]Radboud University, Nijmegen, The Netherlands, [3]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Journal of Cognitive Neuroscience X:Y, pp. 1–17
doi:10.1162/jocn_a_01145

One working hypothesis was that an individual bias in temporal order perception may reflect the spontaneous preference in paying more attention toward one or the other sensory modality. Such intrinsic preference would predict, in agreement with the law of prior entry (Spence & Parise, 2010), that a stimulus being attended to would be perceived as occurring earlier than a simultaneous stimulus being less attended to. Stated differently, an attended stimulus would have priority over another one in the conscious temporal sequencing of sensory information, in agreement with the notion that ongoing neural dynamics endogenously determine the temporal structuring of sensory information (Lashley, 1951). Using a TOJ task, which sensory modality an individual attends to at the time of stimulus presentation should thus predict which stimulus will be perceived first: attending to audition should facilitate auditory processing yielding earlier conscious perception of the sound for near synchrony audiovisual stimuli, and conversely, attending to vision should lead to perceiving the visual event as being first in the pair (Vibell, Klinge, Zampini, Spence, & Nobre, 2007; McDonald, Teder-Salejarvi, Di Russo, & Hillyard, 2005; Zampini, Shore, & Spence, 2005). Several neural indices have been reported supporting this hypothesis: In seminal EEG experiments testing TOJ, the attended stimuli that were more often perceived first also elicited evoked responses with earlier latencies (Vibell et al., 2007) and larger amplitudes (McDonald et al., 2005) than when perceived second. The report of latency shifts in early sensory cortices are consistent with temporal prediction (Stekelenburg & Vroomen, 2007) and parametrical predictions of speech content (van Wassenhove, Grant, & Poeppel, 2005). Additionally, an increase of visual attention has been shown to decrease ongoing prestimulus alpha power in visual cortices (Klimesch, Sauseng, & Hanslmayr, 2007; Sauseng et al., 2005), in turn favoring the detection of an incoming near-threshold stimulus (van Dijk, Schoffelen, Oostenveld, & Jensen, 2008; Hanslmayr et al., 2007; Ergenoglu et al., 2004). Alpha oscillations have been proposed to act as selective filters, which inhibit task-irrelevant stimulus processing in a cyclic or pulsed manner (Sadaghiani & Kleinschmidt, 2016; Klimesch, 2012; Jensen & Mazaheri, 2010). During a TOJ task, a decrease of prestimulus alpha power in auditory cortices together with an increase of alpha power in visual cortices was predicted to index an intrinsic preference in paying more attention to audition associated with an increased likelihood of perceiving the sound first. Conversely, participants whose tendency would be to attend to vision would exhibit a decrease in prestimulus alpha power in the visual cortex associated with an increased likelihood of perceiving the flash first (Figure 1A).

Second, an individual's simultaneity threshold (or Point of Subjective Simultaneity [PSS]) could reflect an individual's bias toward one or the other sensory modality, irrespective of the participant's attentional state (Grabot & van Wassenhove, 2017). For example, one

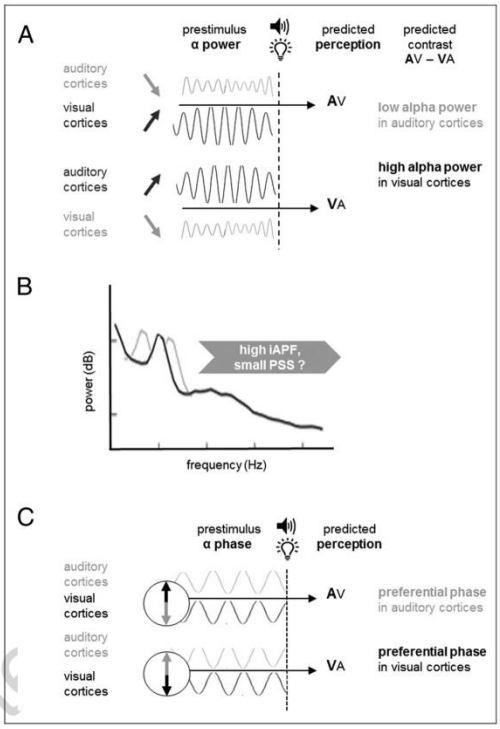


Figure 1. Working hypotheses for prestimulus alpha oscillations. (A) Prestimulus alpha power fluctuations and attention. A decrease of prestimulus alpha power in one sensory modality predicts increased attention to this sensory modality, hence a higher likelihood to detect and report the stimulus as being first. A decrease of prestimulus alpha power in auditory cortices combined with an increase alpha power in visual cortices would thus be predicted when perceiving the sound before the flash; conversely, an increase of prestimulus alpha power in auditory cortices combined with a decrease of prestimulus alpha power in visual cortices would be predicted when perceiving the flash before the sound. (B) iAPF. Following recent reports of correlations between timing ability and higher alpha peak frequency (Cecere et al., 2015; Samaha & Postle, 2015), the higher the individual peak frequency, the finer the discriminability between two sensory events, and the smaller the individual's PSS could be. (C) Phase of prestimulus alpha. Following the hypothesis that the phase of ongoing oscillations indexes excitability, two distinct phases may index the detectability of the stimuli in each sensory cortices resulting in differences of perceived audiovisual temporal order. Specifically, the sensory cortices showing the highest excitability in the prestimulus period should also predict seeing the stimulus in that sensory modality first.

seminal study reported systematic differences between early visual and auditory evoked responses as a function of individuals who may have preferences toward one or the other sensory modality (Giard & Peronnet, 1999). On the basis of the observations that individuals show robust differences in alpha peak frequency (Haegens, Cousijn, Wallis, Harrison, & Nobre, 2014) and that an individual

alpha peak frequency (iAPF) may be predictive of an individual's visual temporal resolution (Cecere, Rees, & Romei, 2015; Samaha & Postle, 2015), we tested whether iAPF could also be a good predictor of subsequent temporal order perception across sensory modalities (Figure 1B).

In a third working hypothesis, we focused on the phase of ongoing neuronal oscillations, in particular alpha and theta oscillations, as predictors of participants' temporal order perception. A growing body of evidence has suggested that the phase of ongoing alpha oscillations predicts the detectability of a sensory event (Mathewson, Fabiani, Gratton, Beck, & Lleras, 2010; Busch, Dubois, & VanRullen, 2009; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009): A stimulus presented at the highest excitability phase of neural oscillations is more likely to be detected and missed otherwise (Schroeder & Lakatos, 2009). The impact of oscillatory phase differences in timing and perception has also been reported across sensory modalities (Thorne & Debener, 2014; Kösem & van Wassenhove, 2012; Thorne, De Vos, Viola, & Debener, 2011). When presented with temporally regular stimuli, cortical oscillations entrain to the input rhythm (Rees, Green, & Kay, 1986; Regan, 1966) and the phase of low-frequency neuronal oscillations has been shown to reflect the temporal expectancy or predictability of event timing (Stefanics et al., 2010). As changes in the phase of entrained neuronal oscillations were previously shown to linearly predict an individual's changes in perceived audiovisual temporal order (Kösem et al., 2014), one hypothesis was that the phase diversity in spontaneous neuronal fluctuations and in the absence of sensory entrainment may predict conscious temporal order. On the basis of the cortical temporal framing hypothesis (Gho & Varela, 1988; Varela, Toro, John, & Schwartz, 1981), the duty cycle of alpha and theta neural oscillations have been hypothesized to code for the temporal sequencing of sensory events (Lisman & Jensen, 2013). The alpha phase, under attentional control (Samaha, Bauer, Cimaroli, & Postle, 2015), has specifically been hypothesized to engage in the prioritization of information in time (Bonnefond & Jensen, 2012; Jensen, Bonnefond, & VanRullen, 2012), thus predicting possible differences in prestimulus alpha phase as a function of perceived order (Figure 1C).

METHODS

Participants

Fourteen participants (6 women, mean age = 24.9 +/− 4.1 years) took part in the study. All had normal or corrected-to-normal vision and normal hearing and were naive as to the purpose of the study. Each participant provided a written informed consent in accordance with the Declaration of Helsinki (2008) and the Ethics Committee on Human Research at NeuroSpin (Gif-sur-Yvette, France). Two participants were discarded from the analysis: One participant showed excessive head movements during MEG recordings, and a second participant did not perform the task properly preventing the assessment of perceptual thresholds. Hence, a total of 12 participants (four women, mean age = 24.8 ± 4.4 years) were analyzed.

Stimuli

The experiment was written in MATLAB 2014a (The MathWorks, Inc., Natick, MA) with the PsychToolbox (version 3.0.11). Visual stimuli consisted of a photodiode turning on for 30 msec in a square wave manner and placed 90 cm away from the participants seated under the MEG dewar. Auditory stimuli consisted of 30 msec sinusoid signals (including 5 msec fade-in and fade-out, with a fundamental frequency of 2 kHz). Sounds were presented binaurally through Etymotic earphones (Etymotic Research, Inc.).

Procedure

The experiment was run in three consecutive steps: 5 min familiarization with the task, followed by 10 min of psychometric assessment, and 40 min of MEG recordings. In all three steps, participants performed a TOJ task on pairs of audiovisual stimuli (Figure 2A): In all TOJ thereafter, a trial consisted of a pair of asynchronous audiovisual stimuli and participants had to indicate by a button press which of the auditory or the visual stimulus was presented first. During familiarization, participants completed 20 TOJ trials to get familiarized with the task and with the motor response mapping. Participants then performed the TOJ task from which we established the psychometric curve and derive the audiovisual SOA used during the MEG recordings in the third part of the experiment (Figure 2B). To establish participants' psychometric curve, 11 SOAs were used: ±300, 240, 180, 120, 60, and 0 msec. A negative (positive) SOA corresponds to auditory leading (lagging) visual stimuli. Each SOA was presented 10 times in random order within and across individuals. Each intertrial interval (ITI) was computed from a standard uniform distribution ranging from 1 to 2 sec. The resulting individuals' psychometric curves were individually fitted with a logistic function, and three parameters were extracted from the fit: the PSS, which corresponds to the SOA when participants report perceiving 50% of "visual stimulus first," and the two Just-Noticeable Differences (JND_1 and JND_2) corresponding to 25% and 75% of "visual stimulus perceived first" responses, respectively. These three individually tailored SOAs were used for the MEG recordings.

The third and last part of the experiment (~40 min) consisted in three MEG acquisition blocks. In each block, the three tested AV asynchronies corresponded to the SOAs obtained for the individual's PSS, JND_1, and JND_2. Each audiovisual asynchrony was presented 30–40 times

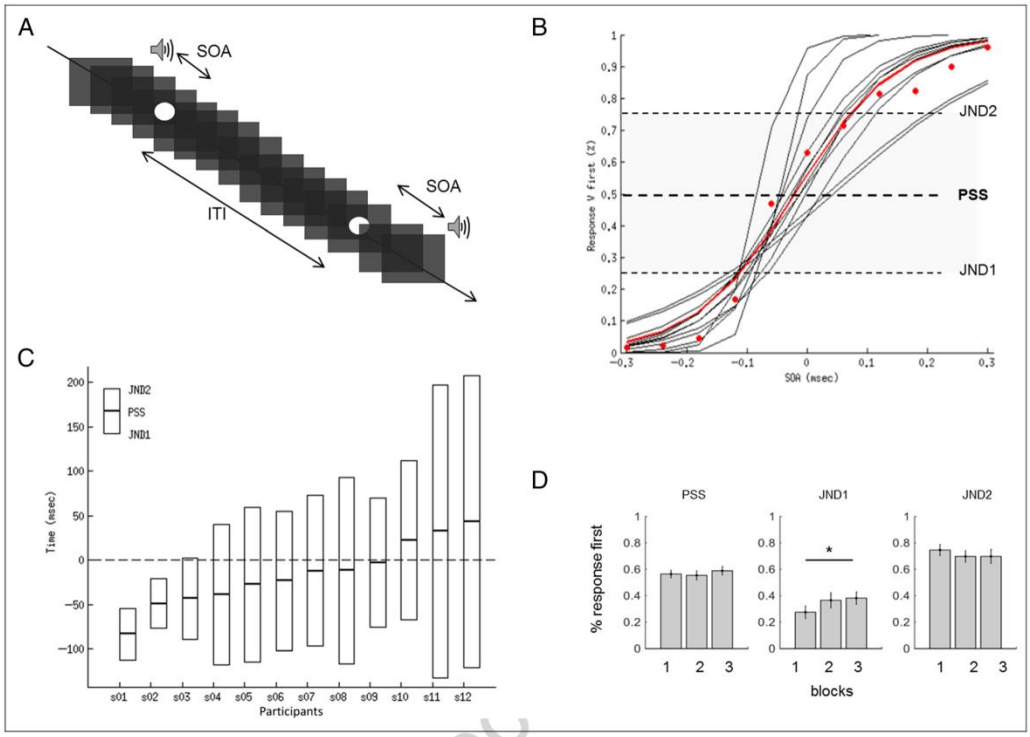


Figure 2. Experimental design and psychophysical results. (A) The experiment consisted in presenting a pair of desynchronized audiovisual stimuli to participants, who performed a TOJ task, that is, judged which of the sound or the visual event came first. A first psychophysics experiment used SOAs selected among 11 values ranging from −300 msec (sound leading) to +300 msec (flash leading) to derive an individual's PSS. ITIs ranged from 1 to 2 sec. During MEG acquisition, only three SOAs were tested tailored to the individual's performance (PSS and JNDs). During MEG, ITIs ranged from 3 to 5 sec. (B) Psychometric curves obtained in the first part of the experiment. The percentage of visual first responses are plotted as a function of audiovisual SOA (msec) for each individual (black) and for the grand average (red). (C) Individual PSS and JND values. A majority of participants showed a negative PSS value indicating that participants required the sound to lead the visual event to perceive them as being simultaneous. PSS and JND values were used as audiovisual SOA in the MEG experiment. (D) Grand-averaged PSS and JND observed during the MEG acquisition in Blocks 1, 2, and 3. No significant changes in PSS (or JND_2) were observed, but a significant increase of JND_1 was found from Blocks 1 to 3.

per block (four and eight participants, respectively). Thus, 90–120 trials per condition was tested during the entire MEG session. The ITIs were selected from a standard uniform distribution ranging from 3 to 5 sec. Long ITIs were chosen for two main reasons: Long delays prevented oscillatory entrainment to the stimuli and insured sufficient time periods for time–frequency analysis in the low-frequency regions (e.g., a minimum of three cycles was available to analyze 1 Hz oscillatory activity). JND_1 and JND_2 conditions were added to balance task difficulty and prevent decreasing attention from the participant. Specifically, JND_1 and JND_2 stimuli could provide clear sensory evidence that could confirm or counteract a participant's intrinsic temporal order bias the PSS condition, by definition, could not. Considering that behavioral outcomes served as our main criterion to classify and contrast the data, preserving only the PSS data meant that the ambiguity in sensory evidence was most likely preserving the initial bias in participants' decision-making. In other words, this condition was the most likely to provide the most legitimate quantification of intrinsic bias during the prestimulus period. Hence, only one contrast reported in Figure 4 used all combined conditions (i.e., JND_1, JND_2, and PSS). Otherwise, data collected for JND_1 and JND_2 were not used for the MEG analysis, which we specifically focused on the PSS condition yielding maximal perceptual ambiguity on a per individual basis.

MEG Data Acquisition

Brain activity was recorded in a magnetically shielded room using the whole-head Elekta Neuromag Vector View 306 MEG system (Elekta Neuromag, Helsinki, Finland) equipped with 102 triple-sensors elements (two

orthogonal planar gradiometers, grad1 and grad2 in text, and one magnetometer per sensor location). Participants were seated in upright position, and their head position was measured before each block using four head position coils placed over the frontal and the mastoid areas. The four head position coils and three additional fiducial points (nasion, left and right preauricular areas) were used during digitization to help with coregistration of the individual's anatomical MRI. MEG recordings were sampled at 1 kHz and band-pass filtered between 0.03 and 330 Hz. The EOGs (horizontal and vertical eye movements) and electrocardiograms were recorded simultaneously with MEG. Before each experiment, a ~1 min empty room recording (no participant in the room) was acquired for the computation of the noise covariance matrix.

MEG Data Analysis

Preprocessing

Signal space separation was applied to decrease the impact of external noise on recorded brain signals (Taulu & Simola, 2006). Gradiometers with amplitudes exceeding 4e−13 T/m were set as "bad sensors" and excluded from further analysis. Signal space separation correction, head movement compensation, and bad sensor rejection was done using MaxFilter Software (Elekta Neuromag). PCA using Graph software (Elekta Neuromag) was performed to remove eye blinks and cardiac artifacts (Uusitalo & Ilmoniemi, 1997). On average, 41 ± 9.8 trials (mean ± 1 *SD*) were used in the PSS contrasts comparing each perceptual outcome (AV and VA).

MEG–aMRI Coregistration

MRI was used to provide high-resolution structural image of each individual's brain. The anatomical MRI was recorded using a 3-T Siemens Trio MRI scanner. Parameters of the sequence were voxel size = 1.0 × 1.0 × 1.1 mm, acquisition time = 466 sec; repetition time = 2300 msec, and echo time = 2.98 msec. Volumetric segmentation of participants' anatomical MRI and cortical surface reconstruction was performed with the FreeSurfer software (surfer.nmr.mgh.harvard.edu/). These procedures were used for group analysis with the MNE suite software (Gramfort et al., 2014). The coregistration of the MEG data with the individual's structural MRI was carried out by realigning the digitized fiducial points with MRI slices. Using *mne_analyze* within the MNE suite, digitized fiducial points were aligned manually with the multimodal markers on the automatically extracted scalp of the participant. To insure reliable coregistration, an iterative refinement procedure was used to realign all digitized points with the individual's scalp.

MEG Source Reconstruction

Individual forward solutions for all source locations located on the cortical sheet were computed using a three-layer boundary element model constrained by the individual's anatomical MRI. Cortical surfaces extracted with FreeSurfer were subsampled to 10,242 equally spaced sources on each hemisphere (3.1 mm between sources). The noise covariance matrix for each individual was estimated from the raw empty room MEG recordings preceding the individual's MEG acquisition. The forward solution, the noise covariance, and source covariance matrices were used to calculate the dynamic statistical parametric mapping estimates (Dale et al., 2000). The inverse computation was done using a loose orientation constraint (loose = 0.4, depth = 0.8) on the radial component of the signal. Individuals' current source estimates were registered on the FreeSurfer average brain for surface-based analysis and visualization.

Time–Domain Source Estimates and Cortical Labels

Both sensor (event-related fields) and source space analyses (source estimates) were performed. Here, for consistency across analyses, we will only report source estimated data distinguishing auditory and visual cortices. Evoked responses obtained after grand averaging across all conditions, all responses, and all participants were used as localizers. Source localizers predictably revealed auditory and visual sensory cortices as main contributors to the MEG signals. Labels encompassing these two regions were selected from the FreeSurfer neuroanatomical parcellation (https://surfer.nmr.mgh.harvard.edu/fswiki/CorticalParcellation). All vertices within these cortical regions were used for all subsequent source space analyses and referred to as "auditory cortices" and "visual cortices" (corresponding to the transverse temporal gyrus and occipital in FreeSurfer, respectively). MEG data were down-sampled to 250 Hz and epoched in 2 sec from 1 sec before the onset of the first stimulus to 1 sec following the first stimulus onset (first stimulus was auditory for participants with a negative PSS and visual for participants with a positive PSS). Epochs were zero-phased (two passes) low-pass filtered at 45 Hz. Baseline correction was applied using the first second of the epoch. Trials were sorted as a function of participants' response, and the number of epochs was equalized across conditions using automatized procedure in MNE-Python, which insured that the selection of trials across conditions were as close in time during the acquisition as possible (cf. martinos.org/mne/stable/generated/mne.epochs.equalize_epoch_counts.html?highlight=equalize#mne.epochs.equalize_epoch_counts).

Power Spectra and iAPF

All epochs were combined, and the power spectrum density (PSD) was computed for the signal ranging from −1 sec to −0.1 sec before the onset of the first presented stimulus. Welch fast Fourier transform was used to compute the PSD for a frequency range of 1–45 Hz. The alpha

peak frequency was manually reported for each individual as the frequency corresponding to the peak maximum within a [8–13] Hz frequency window.

Time–Frequency Analysis

Trials were sorted according to participants' percept. The number of trials was equalized in the same manner as performed in the evoked response analysis. Data were also down-sampled to 250 Hz. Contamination from evoked activity by prior stimuli was prevented by the selection of large ITIs (3–5 sec). Single-trial time–frequency analysis was applied from −3.3 to +2 sec poststimulus onset. As the time–frequency hypotheses focused on prestimulus activity, no baseline was applied for the computation of oscillatory activity (Jensen & Mazaheri, 2010; Busch et al., 2009; Hanslmayr et al., 2007). Power and phase-locking values (PLVs) were calculated using Morlet wavelets for six frequency bands of interest defined by the central oscillatory peak frequency and by their cycles: delta (1.4–2.7 Hz, 3 cycles), theta (3.9–7.1 Hz, 3.4 cycles), alpha (7.4–12.6 Hz, 3.8 cycles), low beta (13–19 Hz, 5.3 cycles), high beta (19.5–30.4 Hz, 4.6 cycles), and gamma (32–48 Hz, 5 cycles). The bifurcation index (BI) was computed on the basis of the PLV (Busch et al., 2009) using

$$\mathrm{PLV}(t) = \frac{1}{N} \cdot \sum_{n}^{N} e^{j \cdot \theta(t,n)}$$

where N was the number of trials and $\theta(t,k)$ was the instantaneous phase at time t for trial n.

$$\mathrm{BI} = [\mathrm{PLV_{AV}} - \mathrm{PLV_{tot}}][\mathrm{PLV_{VA}} - \mathrm{PLV_{tot}}]$$

where AV and VA were the two possible audiovisual temporal order percepts, namely perceiving the sound before the visual event (AV) or perceiving the visual event before the sound (VA). As previously described (Busch et al., 2009), a positive BI indicated that the two conditions were more phase-locked than the total trials together and that their phase differed; otherwise, the total PLV would be identical to the PLV of each condition, and the BI would be around zero. A negative BI indicated that the PLV in one condition was larger than the PLV in the other condition. The power and BI quantifications were averaged on the sensory labels previously determined by the independent source localizer.

Statistical Analysis

Differences in prestimulus alpha power fluctuations were computed for each individual in the auditory and visual sensory cortices and correlated with the individual PSS. Pearson correlations were performed on each time point from −500 to 0 msec relative to the onset of the first physically presented stimulus. The false discovery rate (FDR) correction implemented in MNE-Python was used to correct for multiple comparisons across time (Genovese, Lazar, & Nichols, 2002). Only correlation time points associated with corrected p value under .05 are reported. Phase differences were tested using one-sample cluster-based analysis on BI data testing against zero (Straußß, Henry, Scharinger, & Obleser, 2015). This analysis was run separately in the auditory and visual cortices, using MNE-Python (Gramfort et al., 2013). As no significant differences were found between the two hemispheres, statistical analyses were run on the average. The analysis was restricted to the first 500 msec before the onset of the first physically presented stimulus, and multiple comparisons were corrected by cluster permutation. Additionally, we ran a whole-brain spatiotemporal cluster analysis for each frequency band using the same temporal window. The cluster-forming threshold was set to 0.01, which was equivalent to a t threshold of 3.1 in an experimental design using 12 participants. Only clusters associated with a p value inferior to .05 are reported.

RESULTS

Interindividual Variability in Temporal Order Perception

The mean individual PSS was close to veridical simultaneity with a value of 16 ± 35 msec (*SD*). A large interindividual variability was found consistent with previous reports (Grabot & van Wassenhove, 2017; Ipser et al., 2017; Freeman et al., 2013; Stone et al., 2001). Individual PSSs ranged from −81 msec (sound leading the flash) to +43 msec (light leading the sound). For 9 of 12 participants, the sound needed to be shown before the visual stimulus to be perceived as simultaneous thus yielding to a negative PSS. The mean JND1 and JND2 were −102 ± 20 msec and 70 ± 75 msec, respectively (Figure 2C; Table 1). During the MEG acquisition, the percentage of responses "I perceived the visual stimulus first" resulted in 59% ± 7% (Figure 2D; Table 1), suggesting that changes in audiovisual synchrony perception had taken place in the course of the experiment: In particular, the more negative the initial PSS was (i.e., the more the participant needed the sound to precede the visual stimulus to be perceived as simultaneous), the more deviation toward perceiving the visual event first was found during the MEG session. Seminal recalibration effects in audiovisual synchrony judgments (Fujisaki et al., 2004; Vroomen et al., 2004) suggest that slow perceptual drifts may have occurred during the experiment on a per individual basis. To test this, we looked at the evolution of participants' audiovisual temporal order perception, separately for each MEG block, and we expected that the deviation from the predicted 50% threshold would increase over time. However, there was no significant difference between blocks excepted for a 10.7% increase of $\mathrm{JND_1}$ between Blocks 1 and 3 (Figure 2D; PSS: $t = 0.805$, $p = .439$, 95% CI = [−0.039, 0.083]; $\mathrm{JND_1}$: $t = 2.944$, $p = .015$, 95% CI = [0.024, 0.170]; $\mathrm{JND_2}$: $t = -1.504$, $p = .163$, 95%

Table 1. Individual Behavioral Data before and during MEG

				PSS	*JND_1*	*JND_2*
Participants	*PSS (msec)*	*JND_1 (msec)*	*JND_2 (msec)*	*% "V First" Block Blocks 1-2-3 (Mean)*		
S01	−39	−118	40	60-80-70 (70)	10-20-27 (19)	83-80-90 (84)
S02	−23	−101	55	53-67 (60)	17-13 (15)	73-83 (78)
S03	−27	−115	60	57-50-70 (59)	20-27-32 (27)	90-90-82 (88)
S04	−49	−77	−21	50-52-55 (53)	35-50-42 (43)	65-62-50 (59)
S05	−83	−112	−54	77-60-82 (72)	52-72-70 (65)	87-80-90 (86)
S06	−3	−75	70	37-52-52 (48)	35-40-40 (38)	77-70-75 (74)
S07	−43	−89	3	62-50-57 (62)	35-55-45 (40)	72-60-62 (61)
S08	−12	−97	73	60-62-67 (63)	0-7-7 (5)	70-70-62 (67)
S09	−12	−116	93	60-57-50 (56)	25-45-50 (40)	80-67-65 (71)
S10	43	−121	208	63-30-50 (48)	23-40-43 (35)	63-43-40 (49)
S11	23	−67	112	52-60-52 (55)	17-10-17 (15)	95-87-100 (94)
S12	33	−132	198	42-42-42 (42)	60-57-45 (54)	37-42-50 (43)

The pre-MEG scanning test provide the PSS, JND_1, and JND_2 (left-most columns). The MEG acquisition consisted in three acquisition blocks (1, 2, and 3). Responses « I perceived the flash first » ("V first") were quantified for each condition and each block. The mean across the three blocks is provided in parentheses. Participant S02 did not complete the last block due to technical problems during the MEG acquisition. Figure 2D additionally reports the mean PSS, JND_1, and JND_2 in the MEG blocks.

CI = [−0.122, 0.024]), suggesting that either audiovisual temporal recalibration is a rapid transient process (Van der Burg, Orchard-Mills, & Alais, 2015) or a very slow hysteretic process undetectable over a short time period of time (Martin, Kösem, & van Wassenhove, 2015). Nevertheless, this side observation is consistent with the notion that the PSS is a "snapshot" of a participant's perceptual state, which may be related to the context in which it is measured (Van Eijk, Kohlrausch, Juola, & van de Par, 2008), although intraindividual variability of the PSS in adults is most likely robustly anchored in a stable intrinsic bias (Grabot & van Wassenhove, 2017). The selected SOAs, based on the initial participant's PSS, remained largely ambiguous for the participants, and great care was taken to equalize the trial count in all subsequent MEG contrasts notably in source estimates.

Prestimulus Alpha Fluctuations and iAPF

One working hypothesis was that the differences in prestimulus alpha power could account for the perception of temporal order. First, we ensured that spontaneous alpha oscillations were readily seen in the prestimulus period. Figure 3 illustrates the grand-averaged PSD topographically and for one sensor: The alpha peak power (~10 Hz) could readily be seen. The standard deviation of the PSD was also highest at ~10 Hz, suggesting that the power of spontaneous alpha oscillations showed highest variability across participants and spectral responses in agreement with prior reports (Haegens et al., 2014). Almost all participants displayed a clear alpha peak frequency (Figure 3B), but no significant correlations were found between an individual's alpha peak frequency and an individual's audiovisual PSS (Figure 3C; $r = .11$, $p > .25$). To insure that this result was not due to the low-frequency resolution in the default PSD computations, the PSD was recomputed between −300 and −100 msec with a finer frequency resolution of 0.34 Hz. Despite finer frequency resolution, the correlation between iAPF and PSS remained nonsignificant ($r = .30$, $p = .38$). The correlation between iAPF and JND (as defined by (JND2–JND1)) was also tested and did not yield any significant effect (Figure 3D; $r = -.24$, $p = .47$).

Prestimulus Alpha Power Fluctuations and Group-average Performance

We first combined all conditions (PSS, JND_1, JND_2) and computed the differences of prestimulus alpha power between correct and incorrect trials irrespective of temporal order. A significant decrease in prestimulus alpha power was found when participants were correct. A cluster-based permutation t test contrasting 500 msec prestimulus alpha preceding the presentation of the first stimulus was performed (Figure 4). Three significant clusters were found in sensor space: one widespread cluster including 52 magnetometers for the full 500 msec ($p = .026$) and two local clusters implicating in gradiometers (15 gradiometers, [−456, −12 msec], $p = .023$, and 14 gradiometers,

[−500, 0 msec], $p = .048$). Figure 4A reports the alpha power time course observe across the mean of significant magnetometers for correct (cyan) and incorrect (pink) performance, irrespective of veridical temporal order. The same analysis was carried out in source-reconstructed data and revealed a significant cluster ($p = .003$) ranging from −304 to 0 msec. The cluster was mostly located in the right supramarginal gyrus and auditory regions.

Using performance (correct/incorrect) in this analysis could not dissociate between trials in which audiovisual order confirmed or refuted the a priori bias of an individual's toward one or the other temporal order. Specifically, only stimuli with the SOA corresponding to that of the individual's PSS could by definition preserve the ambiguity in sensory evidence and hence leave unchanged the bias of prestimulus activity; to the contrary, JND_1 and JND_2 provided evidence that could either validate or invalidate the prestimulus bias. Hence, in this analysis, both trials with informative evidence and trials with uncertainty were intermixed leaving uncertain to

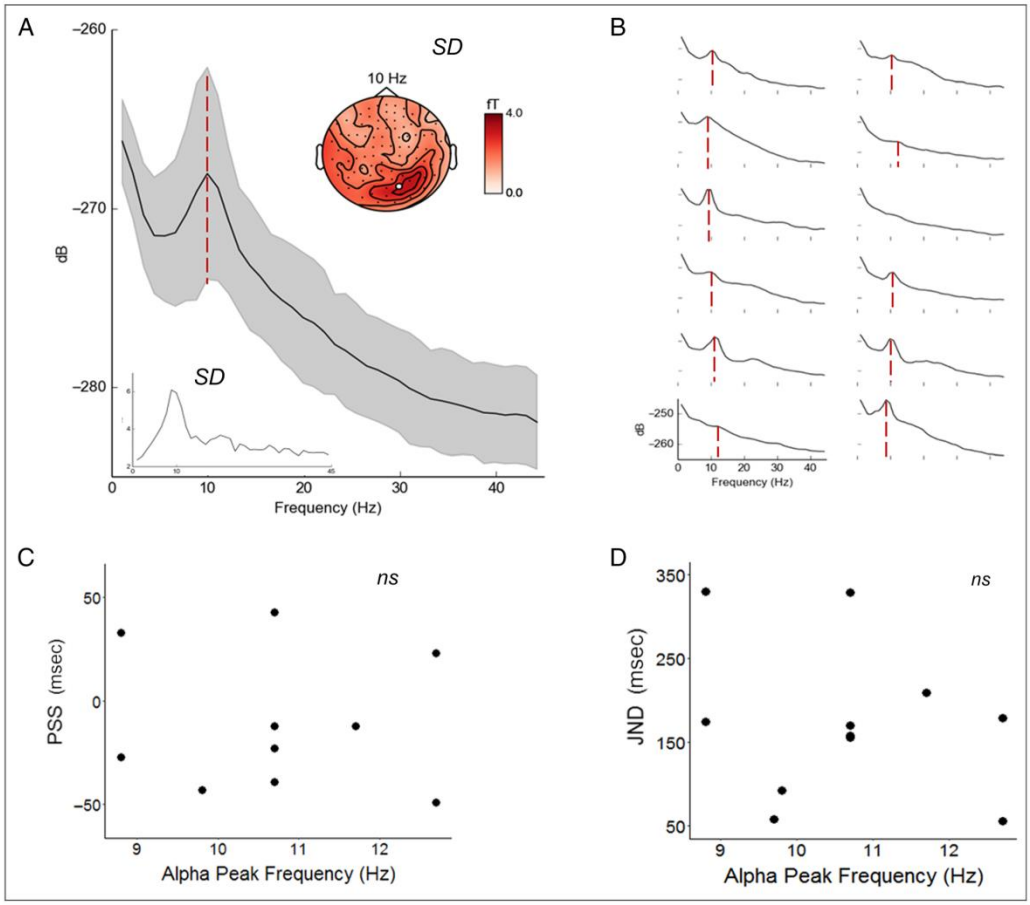


Figure 3. Inter-iAPF variability. (A) For illustration purposes, the group-average PSD is plotted for one representative magnetometer (white disc). The PSD amplitude are in decibels (dB) and calculated for all trials independently of the individual's percept, from −1000 to −100 msec before the onset of the first presented stimulus. The standard deviation (*SD*) of the PSD across all sensors (shaded gray) showed a maximal peak at ~10 Hz (alpha peak frequency). This is also depicted in the bottom inset. The topography (top inset) illustrates the spatial distribution of the standard deviation (*SD*) of the PSD over all sensors at 10 Hz. (B) The PSD averaged across all sensors is displayed for each individual. (C) The alpha peak frequency (red line in B) was extracted for each individual and plotted against the individual's PSS. The correlation between iAPF and PSS was not significant ($r = .06, p = .844$). Increasing the frequency resolution (0.34 Hz) on a larger prestimulus temporal interval (−3000 to −100 msec) did not improve the correlation ($r = -.031, p = .92$; plot not shown). (D) The correlation between iAPF and JND (defined as JND_2–JND_1) was not significant ($r = -.24, p = .47$).

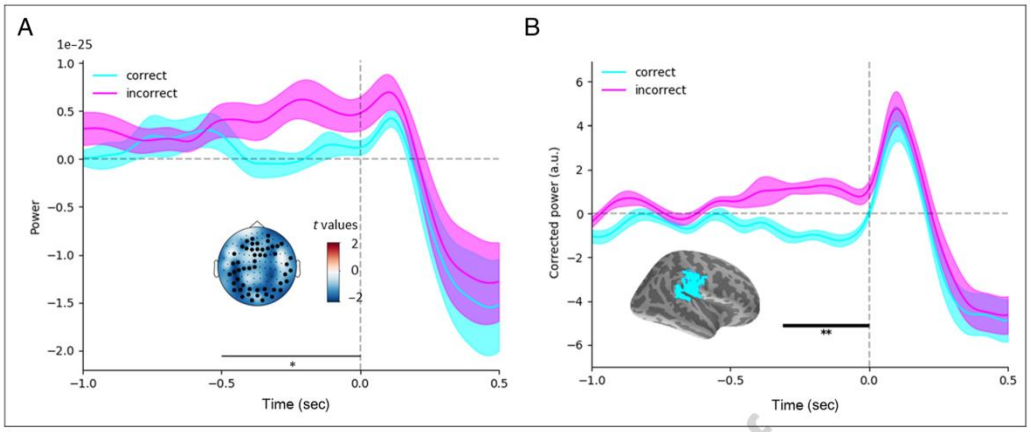


Figure 4. Increased prestimulus alpha power predicts correct performance irrespective of temporal order. In this analysis, all conditions (JND_1, JND_2, PSS) were combined to assess the relationship between prestimulus alpha power changes and performance (correct in blue; incorrect in pink). (A) In sensor space, a significant decrease in prestimulus alpha power was found when the subsequent trial was correctly perceived. A cluster-based permutation *t* test run on each sensor type on the 500 msec prestimulus period (horizontal dashed line) revealed three significant clusters: one widespread cluster including 52 magnetometers ([−500, 0 msec], $p = .026$; topographic inset with *t* values) and two local clusters (15 gradiometers, −456 to −12 msec, $p = .023$, and 14 gradiometers, −500 to 0 msec, $p = .048$). The plotted time course was obtained by averaging the significant cluster of magnetometers highlighted in the inset topographic map. (B) The same analysis was performed on source-reconstructed data and revealed a significant cluster (−304 to 0 msec, $p = .003$) located in the right supramarginal gyrus regions (inset). For illustration, alpha power was corrected by subtracting the mean of the two contrasted performances to the whole epoch. The shaded areas of the time course represents two *SEMs*.

which extent prestimulus alpha power predicts perceptual bias or some form of preparation toward temporal order decision. To disambiguate this issue and specifically focus on the possibility of a temporal order bias in prestimulus alpha, only trials testing the SOA corresponding to individuals' PSS used for subsequent analyses allowing to preserve maximal perceptual ambiguity on a per individual basis.

Interplay between Prestimulus Alpha Power Fluctuations and Individual Biases

We computed the differences of prestimulus alpha power between the two perceptions of temporal order (i.e., perceiving the sound first minus perceiving the flash first [AV – VA]) on a per individual basis and independently in auditory and visual sensory cortices. A strong trend toward a lower power of spontaneous alpha was observed when perceiving AV as compared with perceiving VA in several posterior and parietal cortical regions (Figure 5A). However, and on average, no significant effects of prestimulus alpha power were found as a function of perceived temporal order in the selected sensory cortices or using a whole-brain analysis corrected for multiple comparisons.

However, when looking at the prestimulus time–frequency differences as a function of perceived temporal order for the same physical stimulus, systematic differences in the alpha power were not occurring in the same direction across individuals. As illustrated in Figure 5B, when one individual showed less prestimulus alpha power in auditory cortices when reporting perceiving AV as compared with when perceiving VA (top), another participant would show the opposite pattern with a stronger prestimulus alpha power when reporting perceiving AV as compared with VA (bottom). Considering the observed interindividual variability in PSS (Figure 2B and C), we asked whether the interindividual variability in prestimulus alpha power differences was indicative of an individual's bias toward perceiving one or the other temporal order in the audiovisual sequence.

To test this, we first correlated each individual's PSS with the individual's prestimulus alpha power difference separately for visual and auditory cortices (Figure 5C). A positive correlation between the difference of alpha power as a function of perceived order (AV − VA) and the individual's PSS was found in auditory cortices between −420 and −192 msec before the first stimulus onset (Figure 5C, left inset). This result indicated that a decrease of prestimulus alpha power in auditory cortex could predict perceiving the sound first (AV) if and only if the individual was biased toward perceiving visual first (negative PSS, requiring the presentation of the sound first for audiovisual simultaneity perception). Conversely, a decrease in prestimulus alpha power in auditory

cortices predicted perceiving the visual event first if and only if the participant was biased toward perceiving auditory first (positive PSS, requiring the presentation of a visual event first for audiovisual simultaneity perception). In other words, a decrease in prestimulus alpha power in the contrast (AV − VA < 0) indicated that the participant would correctly report the sound as occurring first, whereas an increase (AV − VA > 0) indicated that the participant would correctly report the visual event as first. This effect was not observed in visual cortices.

A simpler alternative for the interpretation of fluctuations in prestimulus alpha power would be that such power changes correlate with the individual's performance. However, we contend that this would be an incorrect interpretation of results reported in Figure 5C. If prestimulus alpha power solely reflected an individual's

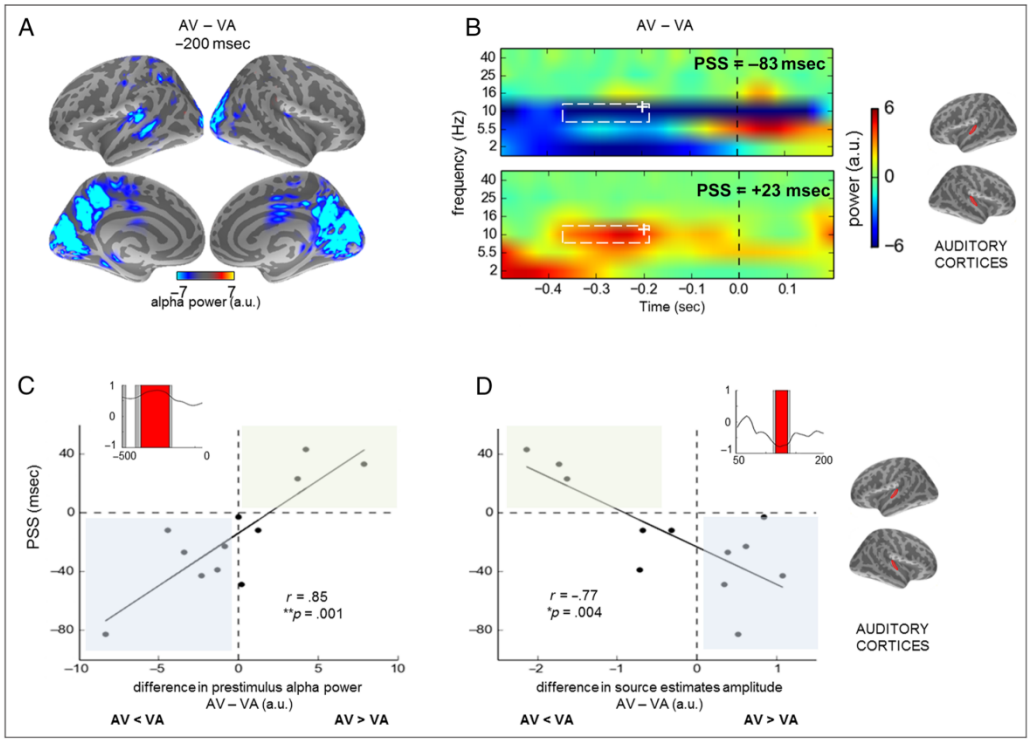


Figure 5. Temporal order perception and prestimulus alpha power. (A) The grand-averaged prestimulus alpha power was contrasted between perceiving auditory first (AV) and perceiving visual first (VA) across all individuals. The snapshot of source-reconstructed brain activity was taken at −200 msec before stimulus onset and corresponds to the white cross (10 Hz) indicated in B. The left and right hemispheres are on the left and right panels, respectively. The external and medial views of source reconstructed brain activity are on the top and bottom panels, respectively. Differences in alpha power were mostly seen in sensory and parietal cortices but did not lead to significant differences at the group level when corrected for multiple comparisons. (B) Example of time–frequency differences for two individuals in auditory cortices. Red indicates higher power when the sound was perceived first (AV); blue indicates higher power when the flash was perceived first (VA). As can be seen on these two time–frequency plots, the sign of the prestimulus alpha power differed across individuals suggesting intersubject variability of prestimulus alpha power as a function of perceived temporal order. (C) Individual PSS as a function of alpha power differences (AV − VA). Each point in the correlation represents an individual. The differences of alpha power between perceiving the sound first (AV) and perceiving the visual event first (VA) were averaged across the right and left auditory cortices. After FDR corrections, latencies showing significant power differences ranged from −420 to −192 msec in the auditory cortices (left). No significant correlations were found in visual cortices (right). These results show a significant correlation between the changes in prestimulus alpha power in auditory cortices and the perceived audiovisual order. Irrespective of the signed value of the power difference or PSS, the larger the power difference, the larger the individual's PSS. (D) The difference of source estimate amplitudes between perceived order (AV − VA) was computed individually from 50 to 200 msec poststimulus onset (time window covering the period of evoked responses) and averaged across hemispheres. A Pearson's correlation was run on each time point between the obtained amplitude difference and the PSS. Gray cluster are significant uncorrected p values of the correlation ($p < .05$), and red clusters are significant FDR-corrected p values. A significant correlation was found in the auditory cortices from 110 to 138 msec. The correlation between differences in source estimates amplitudes and PSS is plotted for this time window. Each point corresponds to a participant.

performance, we would not observe a linear relationship between prestimulus alpha power and individual PSS. Rather, we should observe a categorical pattern with the negative PSS participants (blue) showing negative alpha power in the contrast perceiving AV (correct) versus perceiving VA (incorrect) and the positive PSS participants (green) showing positive alpha power for the same contrast perceiving AV (incorrect) versus perceiving VA (correct). Instead, as seen in Figure 5C, we report a significant linear relationship between alpha power and PSS, suggesting that the magnitude of the prestimulus alpha power changes captures the perceived temporal distance between the auditory and the visual event. To further counteract the simpler explanation of prestimulus alpha power solely reflecting performance, the grand-averaged prestimulus alpha power was computed using all PSS trials, in the same prestimulus window used in analysis reported in Figure 5C (−420 to −192 msec) and in the same cortical label. The correlation between individual performance and prestimulus alpha power was not significant ($r = .27$, $p = .403$).

To investigate further this linear relationship, we looked at the effect of alpha power changes on subsequent auditory evoked responses. Prestimulus alpha power has typically been shown to correlate with the amplitude of the following evoked responses so that increases in alpha power tend to be followed by larger sensory evoked responses (e.g., van Dijk et al., 2008; Sauseng et al., 2005; Ergenoglu et al., 2004). We thus tested whether the amplitude of the source estimates in sensory cortices followed a pattern similar to the one found for prestimulus alpha with PSS. On a per individual basis, we computed the amplitude differences of the auditory current source estimates and correlated the obtained amplitude difference with the individual PSS for each time points between 50 and 200 msec poststimulus onset (Figure 5D). A significant correlation between the amplitude difference in the early auditory source estimates and the individual PSS was found between 110 and 138 msec (Figure 5D; red: FDR-corrected p values). This correlation indicated that larger auditory source estimates were associated with the correct perception of audiovisual temporal order, consistent with the preceding variation of prestimulus alpha power.

Altogether, these results suggested that changes in prestimulus alpha power and evoked source amplitudes were only interpretable insofar as an individual's PSS is known. The interpretation of these results could lend themselves to a prediction of performance on the task but additional observation of the correlations between alpha power differences (Figure 5C) and auditory source amplitude differences (Figure 5D) with individual PSS also showed that the signed magnitude difference in both indices was predictive of an individual's PSS: for instance, the individual's with the most negative PSS also showed the lowest alpha power (and highest auditory source amplitude increase) when reporting perceiving AV as compared with VA.

No Systematic Differences in the Phase of Ongoing Neuronal Oscillations as a Function of Perceived Order

To test the hypothesis of temporal encoding in the phase of neural oscillations, the BI contrasted trials in which participants reported perceiving AV and trials in which participants reported perceiving VA on a per individual basis. Across all frequencies, no positive cluster in the prestimulus period was found. Significant negative clusters were found coinciding with the onset of the first physically presented stimulus. A cluster analysis ran over the full time period of −500 to 1000 msec relative to the onset of the first stimulus revealed significant negative clusters in all frequency bands under 30 Hz at the onset of the evoked responses in both auditory and visual cortices (Figure 6A). A whole-brain analysis on a per frequency basis revealed significant clusters at the latency of the evoked responses in several brain regions; these are illustrated in Figure 6B for theta, alpha, and beta. The significant negative BI meant that the PLV in one condition was stronger than in another condition. This was observed in both auditory and visual cortices. However, a closer look at the systematicity of the BI difference across individuals, for instance, in the alpha band (Figure 6C), revealed no clear distinction as a function of the individual's PSS. The lack of conclusiveness with BI quantifications converges with the results obtained in the differences of evoked source amplitudes showing that the direction of the effects are individually based and need to take into account the individual's temporal order bias.

DISCUSSION

In this MEG study, we report evidence that the degree of synchronization in prestimulus alpha regulates an individual's performance in a TOJ task: For individuals with a negative PSS, prestimulus auditory alpha power was lower when a sound was perceived first than when a visual event was perceived first; conversely, for individuals with a positive PSS, prestimulus auditory alpha power was higher when a sound was perceived first than when a visual event was perceived first. Hence, an individual's perceived temporal order could only be inferred from prestimulus alpha power if and only if the individual's psychological bias (PSS) toward perceiving one or the other sensory modality was known in advance. This result suggests that the interpretability of prestimulus alpha power fluctuations for temporal order perception is a function of an individual's intrinsic bias. Second, we observed that the signed amplitude differences in prestimulus alpha power was correlated with the individual's temporal order bias, further suggesting that prestimulus alpha power does not solely predict performance but rather the extent to which the intrinsic bias must be counteracted to produce a correct response. Third, we found no evidence for individual alpha peak correlations

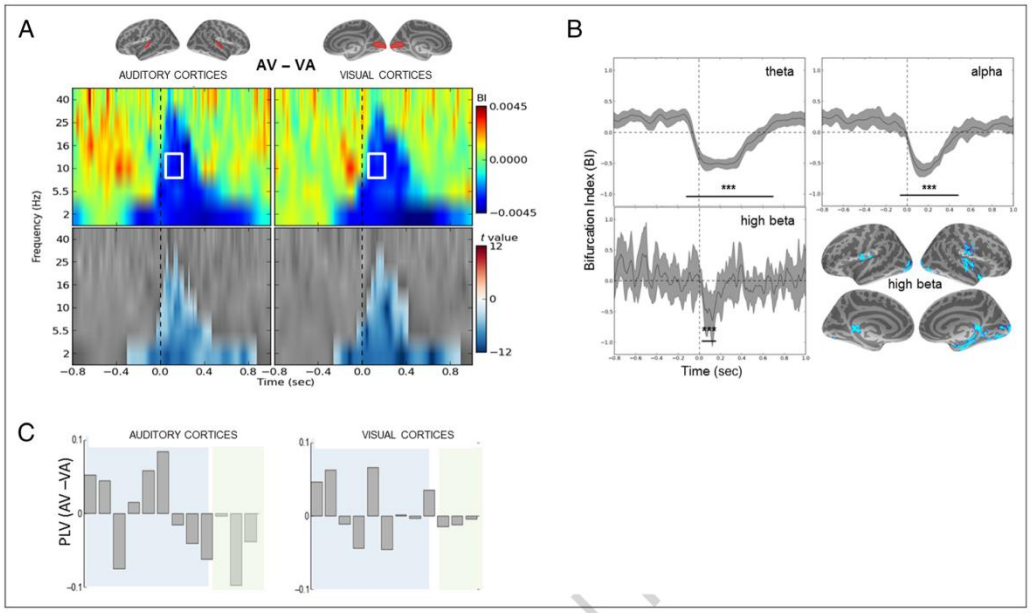


Figure 6. BI and PLVs. (A) BI in auditory (left) and visual (right) cortices contrasting perceiving the sound first and perceiving the visual event first (AV − VA). A significant negative BI was found at the time of the evoked responses (lower panels reporting corrected *t* values). Specifically, the PLVs when perceiving the sound first (AV) were smaller than when perceiving the visual event first (VA). (B) Source estimations of the BI observed across frequency and brain regions for theta, alpha, and high beta (the lower the frequency, the less reliable the latency due to the temporal spread). The peak differences in high-frequency region coincides well with the timing of the early evoked responses. (C) Differences of PLVs on a per individual basis sorted as a function of increasing PSS in the auditory and visual cortices (left and right panels, respectively). The changes in PLV poststimulus onset were not correlated with PSS. This is consistent with the fact that the BI does not take into account the directionality of the PLV differences. Blue shading are individuals with negative PSS, and green shading are individuals with negative PSS (cf. Figure 4). The lack of conclusiveness with BI quantifications is consistent with the results obtained in the differences of evoked source amplitudes showing that the interpretability of the effects need to incorporate the individual's temporal order bias or PSS.

with audiovisual temporal order perception (as indexed by PSS or JND). Fourth, in the absence of sensory entrainment, we found no evidence for a spontaneous preferential phase of ongoing oscillatory activity that would systematically indicate an individual's perceived temporal order. We discuss below the implications of our findings for existing interpretations on the role of spontaneous brain rhythms on the temporal sequencing of (multi)sensory information.

The perception of temporal order across sensory modalities is traditionally seen as being dominated by endogenous attentional orienting considering that the attended stimulus (or sensory modality) is perceived earlier than the unattended stimulus, following Titchener's law of prior entry (Spence & Parise, 2010). In TOJ tasks, manipulating attention has been systematically shown to shift an individual's PSS so that the attended stimulus or sensory modality can be perceived more often as occurring first than the unattended stimulus or sensory modality (Spence, Shore, & Klein, 2001). In the current experimental setup, participants' attention was not explicitly manipulated and was instead predicted to freely fluctuate from one to the other sensory modality. To further increase uncertainty on a per individual basis, the temporal delays between audiovisual events corresponded to the individual's measured perceptual threshold (PSS).

In a first working hypothesis (Figure 1A), we predicted that when participants' attention was biased toward audition (vision), the auditory (visual) event would be most often reported first. One index of attentional fluctuation is the degree of alpha (de)synchronization so that a local decrease in alpha power indexes a decrease of sensory inhibition (Klimesch, 2012; Jensen & Mazaheri, 2010; Sauseng et al., 2005): Previous reports have shown that a decrease of prestimulus alpha power in sensory cortices was correlated with better performance (Mazaheri et al., 2014; Yamagishi, Callan, Anderson, & Kawato, 2008; Hanslmayr et al., 2007) whereas an increase of prestimulus alpha power could yield higher illusory reports (Lange, 2014). Our first set of results (Figures 4 and 7) support the notion that, on average, lower prestimulus alpha power is associated with correct performance irrespective of the actually perceived order.

Additionally, a decrease in prestimulus alpha power was predicted in the sensory cortices corresponding to the first stimulus being correctly reported, namely, for participants with a positive PSS and reporting the visual event as being first, we expected a decrease (increase) of alpha power in visual (auditory) cortices whereas for participants with a negative PSS, reporting the sound as being first would show a decrease (increase) of alpha power in auditory (visual) cortices. In the contrast perceiving (AV − VA), we thus expected consistent decrease of prestimulus alpha power in auditory cortices and increase of alpha power in visual cortices. This is not what we observed.

First, significant changes of alpha power were exclusively observed in auditory cortices. One possibility for the absence of prestimulus alpha power changes in visual cortices in our experiment is the consistency of attentional allocation to the site of visual stimulation (Slagter, Prinssen, Reteig, & Mazaheri, 2016). Another possibility is that sounds were effectively presented first for a majority of participants, thereby biasing the variability in allocation of attention to auditory inputs; future investigations including a larger number of participants with a positive PSS would be needed to see whether the sole implication of auditory cortices is effectively key for event timing (Kösem et al., 2014; Kanai, Lloyd, Bueti, & Walsh, 2011). Alternatively, this observation may suggest that participants did not readily switch or change their attention across sensory modalities and instead adopted a strategy to use auditory information as the main decisional anchor. This would be coherent with the observation that, across individuals, lowest desynchronization in auditory cortices was correlated with being biased toward vision (Figure 7, blue), requiring participants with the largest bias (or negative PSS) to pay more attention to audition and, conversely, those with the largest positive

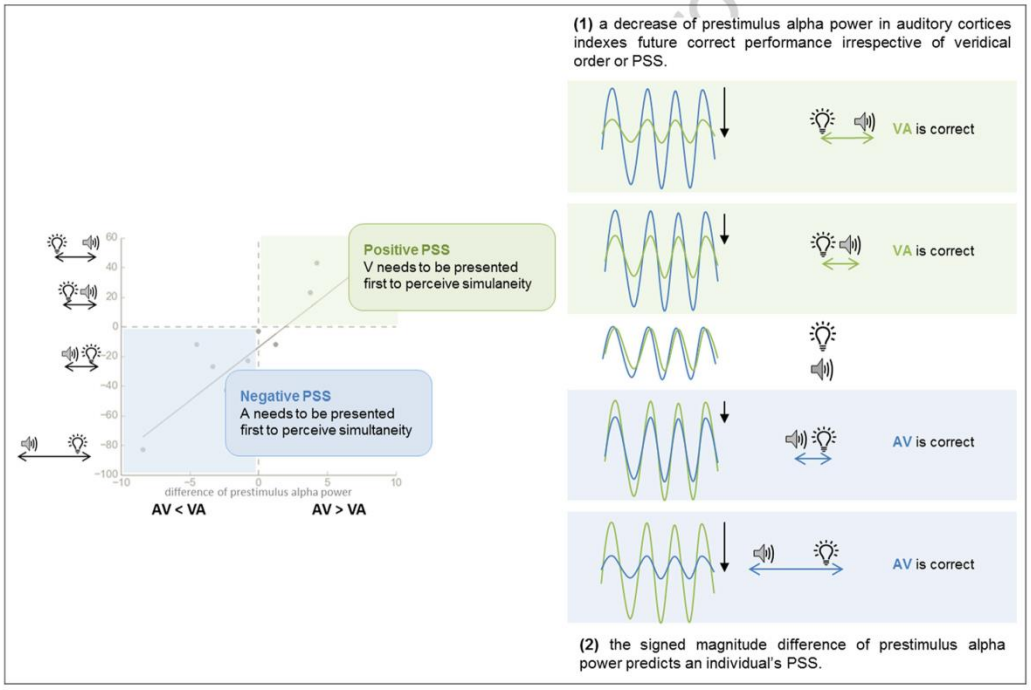


Figure 7. Summary of findings. Using an audiovisual TOJ task whose difficulty was calibrated on each individual's temporal order bias (PSS), we report two main findings. First, the relative changes of prestimulus alpha power in auditory cortices is indicative of an individual's performance (1). Importantly, performance in TOJ specifies the perceived sequence of sensory events. Hence, these results suggest that fluctuations in the power of prestimulus auditory alpha may play an important role in the temporal sequencing of sensory events. Crucially, however, a power increase of alpha was insufficient to predict a participant's perceived temporal order considering that a participant with a negative PSS (blue) showing a relative decrease of alpha power in the contrast (AV − VA) will correctly report perceiving the sound first (i.e., AV percept) whereas a participant with a negative PSS (green) showing the same relative decrease using the same contrast will have perceived the visual event as being first (i.e., VA percept). Additionally, the signed magnitude difference in this contrast was indicative of an individual's PSS so that the largest changes in prestimulus auditory alpha synchronization corresponded to the largest PSS values (2). This strongly suggests that prestimulus alpha power desynchronization does not only reflect general attentional mechanisms but rather endogenous mechanisms compensating for an individual's structural biases.

PSS to pay less attention to audition. For instance, splitting participants according to their preferred sensory modalities, that is, according to whether they performed best in audition or in vision during unimodal presentation has previously been shown to dissociate neural signatures of audiovisual integration (Giard & Peronnet, 1999). In this study, the authors showed, in agreement with the inverse effectiveness principle (Stein & Meredith, 1993), that multisensory integration yielded enhanced early latency cortical responses in the nondominant sensory modality. Here, endogenous modulations of spontaneous alpha synchronization could similarly be seen as counteracting an individual's internal bias (PSS) in temporal order so that the largest relative decrease in alpha power (and associated largest relative increase in the amplitude evoked response) would yield correct performance. Overall, our results are consistent with the role assigned to the power fluctuations of spontaneous alpha oscillations in temporal attention (van Diepen, Cohen, Denys, & Mazaheri, 2015; Klimesch, 2012; Hanslmayr, Gross, Klimesch, & Shapiro, 2011), thereby attentional fluctuations regulate the temporal information flow in sensory regions. Our results converge with the notion that sensory alpha reflects feedback modulatory signals during perceptual analysis (Fries, 2015; van Kerkoerle et al., 2014) and, more specifically, that alpha power may modulate neural excitability and influence perceptual biases (Iemi, Chaumon, Crouzet, & Busch, 2016).

Second, the observed coupling between changes in prestimulus alpha power and subsequent changes in phase-locking and amplitude of the auditory evoked responses was consistent with recent discussions on the sensitivity of phase information as measured with PLV (Canavier, 2015; Ding et al., 2013) and BI (vanRullen, 2016). We found increased variability in phase locking in both sensory cortices when participants reported perceiving the visual stimulus first. Although this could suggest that the observed prestimulus fluctuations and the increase trial-to-trial variability may result from varying modes of engagement in the task (Baird, Smallwood, Lutz, & Schooler, 2014), the systematic changes in the power of spontaneous alpha fluctuations may not solely reflect task engagement and task performance. The range of spontaneous alpha power fluctuations on a per individual basis was proportional to the individual's temporal order bias so that the largest the difference of power between the two possible order perceptions given the same physical stimulation, the largest the individual's PSS (Figure 7). In other words, the less sensitive an individual was to audiovisual delays, the larger the changes in synchronization of the prestimulus alpha power. Additionally, the decrease in prestimulus alpha power occurred when participants reported their nonpreferred sensory modality as occurring first. One tentative working hypothesis is thus that the degree of prestimulus sensory alpha synchronization reflects compensatory mechanisms for the temporal limitations imposed by structural biases. Specifically, lower alpha power in auditory cortex would yield a stronger influence of structural priors, thereby perceived temporal order would reflect the individual's bias whereas higher alpha power would more effectively weight sensory evidence in the decision process. This interpretation would be compatible with the findings that reduced alpha power may not always predict better visual processing but rather enhanced excitability (e.g., Mazaheri et al., 2014; Lange, Oostenveld, & Fries, 2013) in line with inhibitory alpha frameworks (Jensen et al., 2012; Klimesch et al., 2007). The notion that higher prestimulus alpha power may reflect ongoing inhibition of sensory evidence so that structural priors may have a stronger weight in the perceptual decision is also well in line with recent findings (e.g., Iemi et al., 2016; Chaumon & Busch, 2014; van Dijk et al., 2008).

Changes in the phase response entrained by audiovisual stimuli have been shown to follow an individual's temporal order perception (Kösem et al., 2014), suggesting possible endogenous control of the oscillatory phase response in spite of entrainment—that is, nonstationarity (van Wassenhove, 2016). The endogenous control of oscillatory phase responses have previously been reported in temporal expectation and attention tasks (Samaha et al., 2015; Cravo, Rohenkohl, Wyart, & Nobre, 2011), with a direct link between perceived timing and alpha phase (Chakravarthi & Vanrullen, 2012; Gho & Varela, 1988; Varela et al., 1981). A recent study has also provided evidence supporting the gamma dependency on the theta phase in sequence order (Heusser, Poeppel, Ezzyat, & Davachi, 2016). As the relationship between gamma power and the phase of theta and alpha neuronal oscillations may be crucial to determine the order of events, phase dependency of order perception on the phase of ongoing theta or alpha neuronal oscillations was expected in our task (Figure 1C). However, no systematic evidence supported the role of oscillatory phase in audiovisual temporal order, and we discuss below several possible reasons.

First, one important observation is that previous results tested visual perception and implicit timing (Chakravarthi & Vanrullen, 2012; Gho & Varela, 1988; Varela et al., 1981) whereas our study tested conscious temporal order across sensory modalities. Additionally, we specifically focused on prestimulus activity and reported that phase diversity, at the moment of sensory stimulation, did impact PLVs across the spectrum, but in a manner that seemed individually specific. Second, previous results have tested sequence order in memory (Heusser et al., 2016), whereas our study tested the perceptual encoding stage of sequence order, leaving open the possibility that a different strategy may be used during the encoding of sensory sequences than during memory maintenance of these sequences. Third, a recent EEG study reported differences in the phase of prestimulus alpha oscillations using a very similar design as ours, but perhaps importantly, asking participants to report the

simultaneity of audiovisual events (Ikumi, 2016; Ikumi, Torralba, Ruzzoli, & Soto-Faraco, 2015): One possibility for the differences in findings is that alpha phase and power effects in prestimulus oscillatory activity for the same physical stimulation may be strongly driven by task demands. For instance, the resetting of theta cycles have been shown to discriminate between sensory integration and sensory segregation, providing relevant cues on the extent of temporal organization in the brain (Wutz, Muschter, van Koningsbruggen, Weisz, & Melcher, 2016). In simultaneity tasks, the emphasis is put on the integration of multisensory information, whereas in TOJ tasks, the emphasis is put on the segregation of audiovisual events to elicit a conscious representation of temporal sequence. We thus hope that future studies will help provide direct contrasts between these two forms of temporal encoding multisensory information.

Acknowledgments

This work was supported by ERC-YStG-263584 to V. van Wassenhove. We thank members of UNIACT for their help in recruiting and scheduling participants. Preliminary data were presented at SFN Washington DC 2014, OHBM Honolulu 2015, and Rovereto Attention Workshop 2015. We thank two anonymous reviewers for the quality and insightfulness of their comments on earlier versions of the manuscript.

Reprint requests should be sent to Laetitia Grabot, Cognitive Neuroimaging Unit, CEA DRF/I2BM, INSERM, Université Paris-Sud, Université Paris-Saclay, NeuroSpin Center, 91191 Gif/Yvette, France, or via e-mail: laetitia.grabot@gmail.com.

REFERENCES

Baird, B., Smallwood, J., Lutz, A., & Schooler, J. W. (2014). The decoupled mind: Mind-wandering disrupts cortical phase-locking to perceptual events. *Journal of Cognitive Neuroscience, 26,* 2596–2607.

Benevento, A., Fallom, J., Davis, B. J., & Rezak, M. (1977). Auditory-visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. *Experimental Neurology, 57,* 849–872.

Bonnefond, M., & Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Current Biology, 22,* 1969–1974.

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *Journal of Neuroscience, 29,* 7869–7876.

Cecere, R., Rees, G., & Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Current Biology, 25,* 231–235.

Chakravarthi, R., & Vanrullen, R. (2012). Conscious updating is a rhythmic process. *Proceedings of the National Academy of Sciences, U.S.A., 109,* 10599–10604.

Chaumon, M., & Busch, N. A. (2014). Prestimulus neural oscillations inhibit visual perception via modulation of response gain. *Journal of Cognitive Neuroscience, 26,* 2514–2529.

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2011). Endogenous modulation of low frequency oscillations by temporal expectations. *Journal of Neurophysiology, 106,* 2964–2972.

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron, 26,* 55–67.

Ergenoglu, T., Demiralp, T., Bayraktaroglu, Z., Ergen, M., Beydagi, H., & Uresin, Y. (2004). Alpha rhythm of the EEG modulates visual detection performance in humans. *Brain Research, Cognitive Brain Research, 20,* 376–383.

Freeman, E. D., Ipser, A., Palmbaha, A., Paunoiu, D., Brown, P., Lambert, C., et al. (2013). Sight and sound out of synch: Fragmentation and renormalisation of audiovisual integration and subjective timing. *Cortex, 49,* 2875–2887.

Fries, P. (2015). Rhythms for cognition: Communication through coherence. *Neuron, 88,* 220–235.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7,* 773–778.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage, 15,* 870–878.

Gho, M., & Varela, F. J. (1988). A quantitative assessment of the dependency of the visual temporal frame upon the cortical rhythm. *Journal of Physiology (Paris), 83,* 95–101.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11,* 473–490.

Grabot, L., & van Wassenhove, V. (2017). Time order as psychological bias. *Psychological Science*. doi:10.1177/0956797616689369.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., et al. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience, 7*.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., et al. (2014). MNE software for processing MEG and EEG data. *Neuroimage, 86,* 446–460.

Haegens, S., Cousijn, H., Wallis, G., Harrison, P. J., & Nobre, A. C. (2014). Inter- and intra-individual variability in alpha peak frequency. *Neuroimage, 92,* 46–55.

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bäuml, K. H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *Neuroimage, 37,* 1465–1473.

Hanslmayr, S., Gross, J., Klimesch, W., & Shapiro, K. L. (2011). The role of α oscillations in temporal attention. *Brain Research Reviews, 67,* 331–343.

Hanson, J. V. M., Heron, J., & Whitaker, D. (2008). Recalibration of perceived time across sensory modalities. *Experimental Brain Research, 185,* 347–352.

Heusser, A. C., Poeppel, D., Ezzyat, Y., & Davachi, L. (2016). Episodic sequence memory is supported by a theta-gamma phase code. *Nature Neuroscience, 19,* 1374–1380.

Iemi, L., Chaumon, M., Crouzet, S. M., & Busch, N. A. (2016). Spontaneous neural oscillations bias perception by modulating baseline excitability. *Journal of Neuroscience,* 1432-16. doi: https://doi.org/10.1523/JNEUROSCI.1432-16.2016.

Ikumi, N., Torralba, M., Ruzzoli, M., & Soto-Faraco, S. (2015). *Ongoing alpha phase explains perceptual variability in audio-visual synchrony judgments*. Paper presented at the Rovereto Attention Worskhop, Rovereto, Italy.

Ikumi, N. M. (2016). *The effects of cognitive factors on cross-modal synchrony perception*. Barcelona: Universitat Pompeu Fabra.

Ipser, A., Agolli, V., Bajraktari, A., Al-Alawi, F., Djaafara, N., & Freeman, E. D. (2017). Sight and sound persistently out of synch: Stable individual differences in audiovisual synchronisation revealed by implicit measures of lip-voice integration. *Scientific Reports, 7,* 46413.

Jensen, O., Bonnefond, M., & VanRullen, R. (2012). An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends in Cognitive Sciences, 16,* 200–206.
Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience, 4,* 186.
Kanai, R., Lloyd, H., Bueti, D., & Walsh, V. (2011). Modality-independent role of the primary auditory cortex in time estimation. *Experimental Brain Research, 209,* 465–471.
Klimesch, W. (2012). α-Band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences, 16,* 606–617.
Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews, 53,* 63–88.
Kösem, A., Gramfort, A., & van Wassenhove, V. (2014). Encoding of event timing in the phase of neural oscillations. *Neuroimage, 92C,* 274–284.
Kösem, A., & van Wassenhove, V. (2012). Temporal structure in audiovisual sensory selection. *PLoS One, 7,* e40936.
Lange, J., Oostenveld, R., & Fries, P. (2013). Reduced occipital alpha power indexes enhanced excitability rather than improved visual perception. *Journal of Neuroscience, 33,* 3212–3220.
Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: John Wiley & Sons.
Lisman, J. E., & Jensen, O. (2013). The θ-γ neural code. *Neuron, 77,* 1002–1016.
Martin, J.-R., Kösem, A., & van Wassenhove, V. (2015). Hysteresis in audiovisual synchrony perception. *PLoS One, 10,* e0119365.
Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M., & Lleras, A. (2010). Rescuing stimuli from invisibility: Inducing a momentary release from visual masking with pre-target entrainment. *Cognition, 115,* 186–191.
Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: Prestimulus alpha phase predicts visual awareness. *Journal of Neuroscience, 29,* 2725–2732.
Mazaheri, A., van Schouwenburg, M. R., Dimitrijevic, A., Denys, D., Cools, R., & Jensen, O. (2014). Region-specific modulations in oscillatory alpha activity serve to facilitate processing in the visual and auditory modalities. *Neuroimage, 87,* 356–362.
McDonald, J. J., Teder-Salejarvi, W. A., Di Russo, F., & Hillyard, S. A. (2005). Neural basis of auditory-induced shifts in visual time-order perception. *Nature Neuroscience, 8,* 1197–1202.
Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience, 7,* 3215–3229.
Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 364,* 1887–1896.
Pöppel, E., Schill, K., & von Steinbüchel, N. (1990). Sensory integration within temporally neutral systems states: A hypothesis. *Naturwissenschaften, 77,* 89–91.
Rees, A., Green, G. G., & Kay, R. H. (1986). Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hear Res, 23,* 123–133.
Regan, D. (1966). Some characteristics of average steady-state and transient responses evoked by modulated light. *Electroencephalography and Clinical Neurophysiology, 20,* 238–248.
Sadaghiani, S., & Kleinschmidt, A. (2016). Brain networks and α-oscillations: Structural and functional foundations of cognitive control. *Trends in Cognitive Sciences*. doi:10.1016/j.tics.2016.09.004.
Samaha, J., Bauer, P., Cimaroli, S., & Postle, B. R. (2015). Top–down control of the phase of alpha-band oscillations as a mechanism for temporal prediction. *Proceedings of the National Academy of Sciences, U.S.A., 112,* 8439–8444.
Samaha, J., & Postle, B. R. (2015). The speed of alpha-band oscillations predicts the temporal resolution of visual perception. *Current Biology, 25,* 2985–2990.
Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., et al. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. *European Journal of Neuroscience, 22,* 2917–2926.
Schroeder, C., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences, 32,* 9–18.
Slagter, H. A., Prinssen, S., Reteig, L. C., & Mazaheri, A. (2016). Facilitation and inhibition in attention: Functional dissociation of pre-stimulus alpha activity, P1, and N1 components. *Neuroimage, 125,* 25–35.
Spence, C., & Parise, C. (2010). Prior-entry: A review. *Consciousness and Cognition, 19,* 364–379.
Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General, 130,* 799–832.
Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *Journal of Neuroscience, 30,* 13578–13585.
Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.
Stekelenburg, J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *Journal of Cognitive Neuroscience, 19,* 1964–1973.
Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., et al. (2001). When is now? Perception of simultaneity. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 268,* 31.
Strauß, A., Henry, M. J., Scharinger, M., & Obleser, J. (2015). Alpha phase determines successful lexical decision in noise. *Journal of Neuroscience, 35,* 3256–3262.
Taulu, S., & Simola, J. (2006). Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. *Physics in Medicine and Biology, 51,* 1759–1768.
Thorne, J. D., De Vos, M., Viola, F. C., & Debener, S. (2011). Cross-modal phase reset predicts auditory task performance in humans. *Journal of Neuroscience, 31,* 3853–3861.
Thorne, J. D., & Debener, S. (2014). Look now and hear what's coming: On the functional role of cross-modal phase reset. *Hearing Research, 307,* 144–152.
Uusitalo, M. A., & Ilmoniemi, R. J. (1997). Signal-space projection method for separating MEG or EEG into components. *Medical & Biological Engineering & Computing, 35,* 135–140.
Van der Burg, E., Orchard-Mills, E., & Alais, D. (2015). Rapid temporal recalibration is unique to audiovisual stimuli. *Experimental Brain Research, 233,* 53–59.
van Diepen, R. M., Cohen, M. X., Denys, D., & Mazaheri, A. (2015). Attention and temporal expectations modulate power, not phase, of ongoing alpha oscillations. *Journal of Cognitive Neuroscience, 27,* 1573–1586.
van Dijk, H., Schoffelen, J. M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *Journal of Neuroscience, 28,* 1816–1823.

Van Eijk, R. L., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Attention, Perception, & Psychophysics, 70,* 955–968.
van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M. A., Poort, J., van der Togt, C., et al. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences, U.S.A., 111,* 14332–14341.
van Wassenhove, V. (2016). Temporal cognition and neural oscillations. *Current Opinion in Behavioral Sciences, 8,* 124–130.
van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 1181–1186.
van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia, 45,* 598–607.
Varela, F. J., Toro, A., John, E. R., & Schwartz, E. L. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia, 19,* 675–686.
Vibell, J., Klinge, C., Zampini, M., Spence, C., & Nobre, A. C. (2007). Temporal order is coded temporally in the brain: Early event-related potential latency shifts underlying prior entry in a cross-modal temporal order judgment task. *Journal of Cognitive Neuroscience, 19,* 109–120.
Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics, 72,* 871–884.
Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Brain Research, Cognitive Brain Research, 22,* 32–35.
Wutz, A., Muschter, E., van Koningsbruggen, M. G., Weisz, N., & Melcher, D. (2016). Temporal integration windows in neural processing and perception aligned to saccadic eye movements. *Current Biology, 26,* 1659–1668.
Yamagishi, N., Callan, D. E., Anderson, S. J., & Kawato, M. (2008). Attentional changes in pre-stimulus oscillatory activity within early visual cortex are predictive of human visual performance. *Brain Research, 1197,* 115–122.
Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neuroscience Letters, 381,* 217–222.

Chapter 3

Temporal sequences implicitly shape conscious content

Literature review

3.1. Prediction is carried by low-frequency oscillations

The hypothesis of temporal encoding windows suggests that the perception of a target occurring at a specific time point will be influenced both by what, within the encoding window, has occurred before and after the target. Prediction will be briefly introduced along with its neural correlates and we will then focus on postdiction.

3.1.1. Prediction

The brain constantly emits predictions about forthcoming events, and updates these predictions based on their comparison with sensory evidence (Friston, 2005; Brown & Braver, 2005). These top-down modulation biases perceptual inferences in favor of a plausible solution and could be mediated by the synchronization of ongoing neural oscillations (Engel et al., 2001). In both humans and monkeys, feedback processing was related to alpha/beta rhythms while feedforward information could be carried by gamma oscillations (van Kerkoerle et al., 2014; Michalareas et al., 2016).

The brain can also exploit the temporal regularities or temporal relationships between events to ensure an optimal processing (Arnal & Giraud, 2012) as seen when external rhythms entrain neural oscillations (Lakatos et al., 2008). Anticipating when events occur, or temporal prediction/expectation, has been shown to reset the theta-delta entrained oscillations and align to the predicted onset of the stimulus with the most optimal phase (Fig. 3.1A, Stefanics et al., 2011). Even in the absence of entrainment, top-down expectation can regulate the alpha power (Rohenkhol & Nobre, 2011) and phase (Samaha et al., 2015) so that an attended stimulus will arrive during an optimal phase enhancing visual performance.

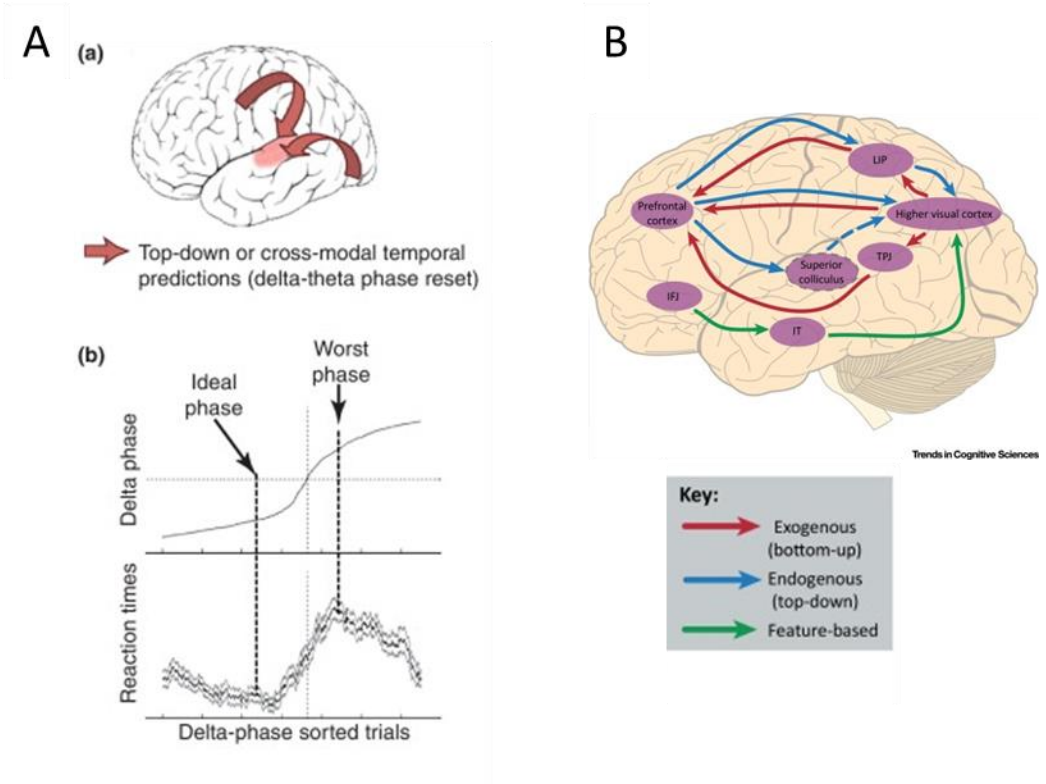


Fig. 3.1 Temporal prediction and selective attention. A. The phase of low-frequency oscillations is reset by top-down temporal expectations in the auditory cortices. The ideal phase within a delta cycle fastens reaction time. **B.** The cartoon depicts selective attention networks with endogenous and exogenous fluxes in blue and red respectively. The broken line indicates that information transmitted by the superior colliculus to the visual cortex transit by the thalamus. Green arrows denote feature-based signals. Abbreviations: IFJ, inferior frontal junction; IT, inferotemporal cortex; LIP, intraparietal sulcus; TPJ, temporoparietal junction. (Adapted from: A: Arnald & Giraud, 2012, B: Mueller et al., 2017)

3.1.2. Prediction differs from endogenous attention

Endogenous attention can also influence perception and differs from prediction in the sense that it does not carry any informational content but rather prioritize processing in a goal-relevant manner. For instance, an abstract cue can indicate in which side of the screen a visual target will appear, which overtly directs spatial attention and leads to increased detectability (Fig. 3.1B). Selective attention has been shown to correlate with increased gamma activity in monkey visual cortices and decreased activity in low-frequency (Fries et al., 2001). The representation of attended location was found in prefrontal cortex, while bottom-up attention involved the parietal cortex (Buschman & Miller, 2007). Prediction and attention modulate neural responses in opposite directions: neural responses evoked by attended stimulus are increased while predicted stimulus triggers reduced neural responses (Arnal & Giraud, 2012).

One recent study orthogonally manipulated attention and predictions (Sherman et al., 2016). Expectations were varied by changing instructions about the probability of target occurrence. It was shown that, independently of attention (Fig. 3.2A), the phase of alpha oscillation at stimulus onset predicted the extent of prior expectations (Sherman et al., 2016, Fig. 3.2B). Importantly, the same alpha phase predicted the absence of a stimulus, whether the expectation was for or against the absence of the target (Fig. 3.2C). The authors proposed that priors are periodically transmitted by the alpha oscillations. An alternatively interpretation could be that a mechanism is regulated the timing of oscillation so that the optimal alpha phase is aligned with the stimulus, depending on *what is expected* (stimulus presence or absence). It also remains unclear whether alpha oscillation is carrying the probability of the prior to influence the final judgment or the prior content itself.

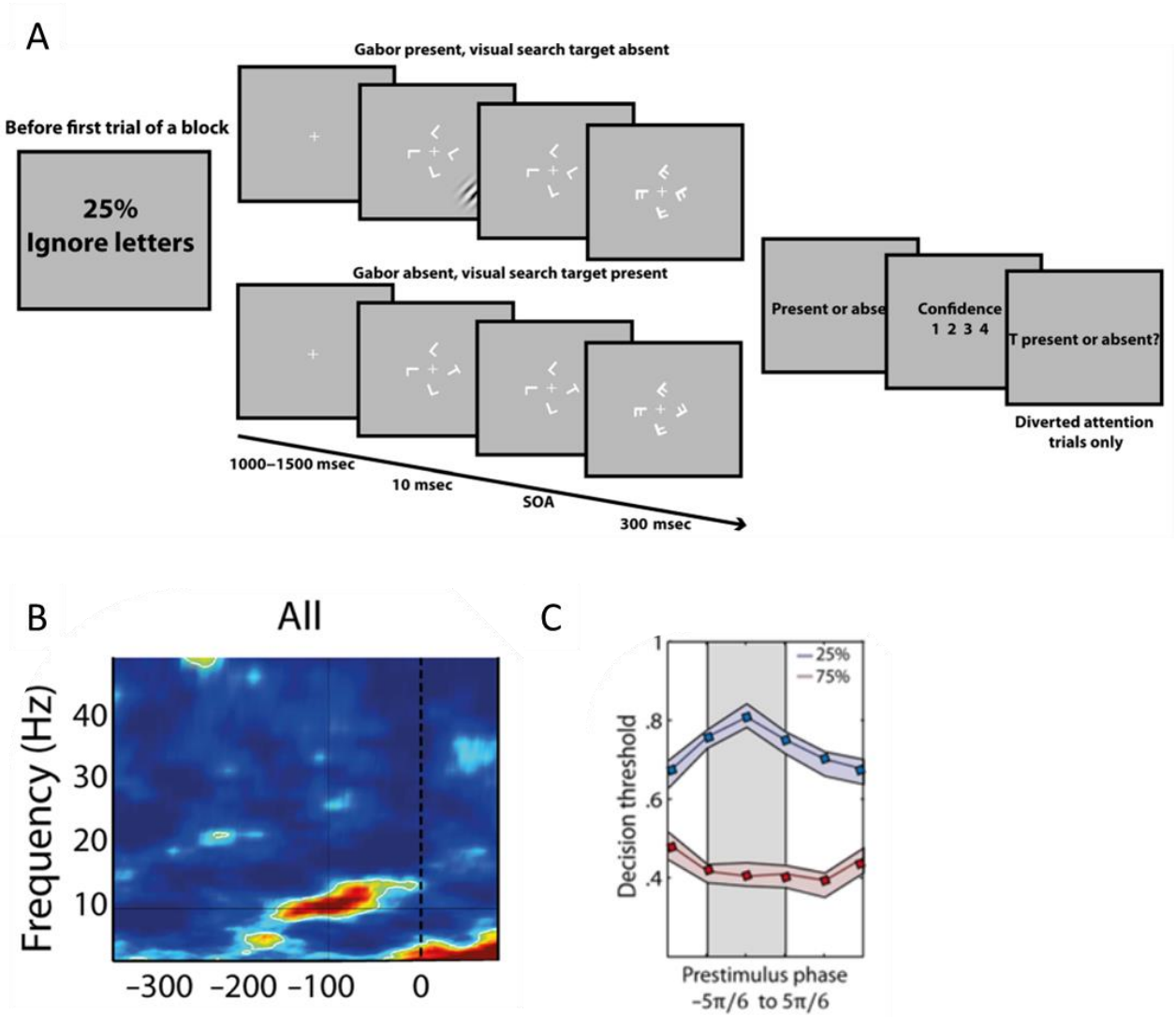


Fig. 3.2 Alpha phase in prior expectations. A. Before the beginning of the trial, indications give the attentional instructions (ignore or attend to the visual search array) and modulate the prior expectation (the Gabor is present in 25% - more likely absent- or in 75% of the trials –more likely present). The participant should report the absence or presence of the Gabor, rate his/her confidence and, in the visual search trial only report the presence of the target letter. **B.** Tim-frequency representation of phase opposition between

yes and no reports for all trials, during prestimulus activity. The vertical dashed line represents the onset of the Gabor. The alpha phase significantly (white outlines) differs between reports. **C.** An optimal alpha phase (grey shading) biases decision toward reporting « no » in the absence-expected condition (25%, in blue), and « yes » in the presence-expected condition (75%, in red). (Adapted from Sherman et al., 2016).

3.2. Defining postdiction

3.2.1. Postdiction

The perception of a stimulus can surprisingly be affected by what happen after its presentation. Postdiction was operationally defined by Shimojo (2014) as the phenomenon in which "a stimulus presented later seems causally to affect the percept of another stimulus presented earlier". Postdiction arguably applies to illusory phenomena like apparent motion, the colored phi phenomenon or the flash-lag effect (Nijhawan, 1994).

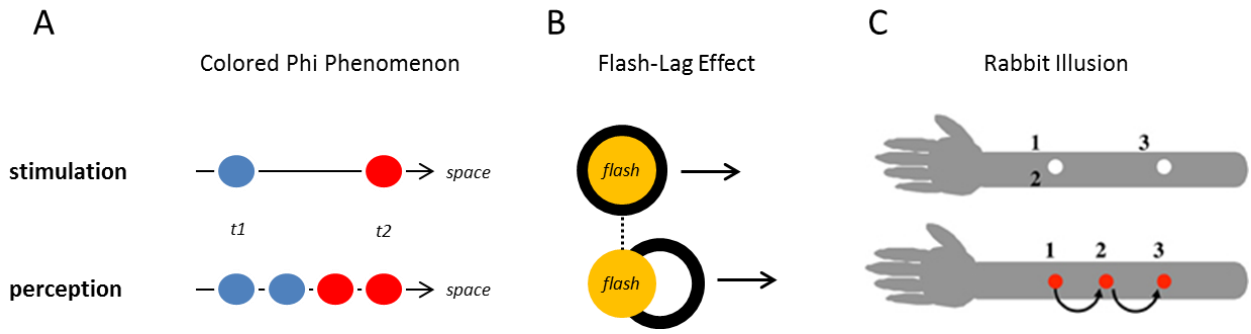


Fig. 3.3 Postdictive phenomena. A. In the colored phi phenomenon, two stimuli presented at an optimal spatio-temporal distance will be perceived as one stimulus moving (apparent motion effect) and changing color from blue to red. **B.** In the flash-lag effect, a moving stimulus (black ring) is perceived as leading a flash whereas both stimuli were simultaneously presented. **C.** In the rabbit illusion, a sequence of three tactile stimuli is presented on the participant's arm, with a regular inter-stimulus interval around 100ms. The second stimulus is tapped at the same spatial location (in white) than the first one, near to the wrist, but is felt closer to the location of the third tap (in red). (Adapted from: C: Shimojo, 2014).

The notion of postdiction historically derived from the illusion of the flash-lag effect, where a moving object is perceived as leading a flash presented at the same position (Fig. 3.3A and B; for a complete review, see Bachmann, 2013). It was first suggested that compensation mechanisms extrapolate the location of the moving stimulus, correcting for the spatial lag due to processing latencies, but fail to do so with the unpredictable flash (Nijhawan, 1994). An alternative hypothesis, the latency difference, proposed that flashed stimuli take more time to be processed than moving objects (Baldo et al., 1995; Whitney & Murakami, 1998). The difference in processing delays was hypothesized to arise

from the attentional reallocation process (Baldo et al., 1995). A seminal set of experiments manipulating the trajectory of the moving object challenged these views and were at the origin of the postdictive account (Eagleman & Sejnowski, 2000). A flash was presented at the same spatial position than a moving ring, and was perceptually mislocalized. The illusion depended not on the pre- but only on the post-flash trajectory of the object, the computation of the trajectory being reset by the flash onset (Fig. 3.4). One criticism of postdiction was that the flash reset predicts a change in perceived speed, which fails to be empirically observed (Nijhawan, 2002).

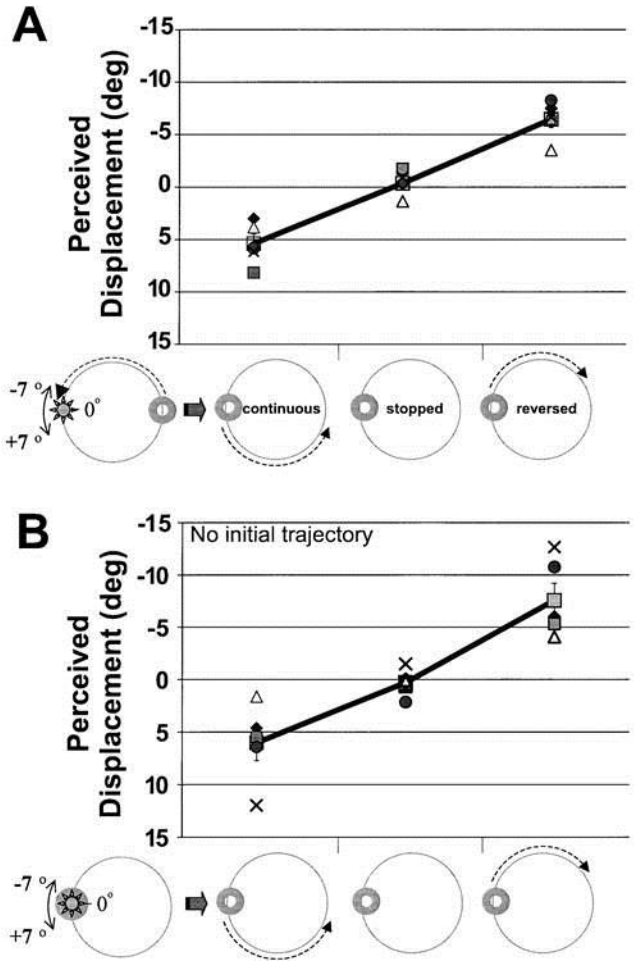


Fig. 3.4 Seminal study on postdiction. A. A flash was presented at the same spatial position than a moving ring, and participants were asked to report whether they perceived the flash below or under the center of the ring. The ring continues its trajectory (left condition), stops (middle), or reverses its trajectory (right). **B.** In the flash-initiated condition, the moving ring started at the spatial location of the flash (left), did not move (middle) or moved forward (right). These experiences showed that the mislocalization depends on the post-flash trajectory of the object. (Adapted from Eagleman & Sejnowski, 2000).

The flash-lag effect probably arises from a combination of different mechanisms, which could unify these findings (see also for more exhaustive reviews, Nijhawan, 2002; Hubbard, 2014; Khoei et al., 2017). A very recent study proposes a Bayesian model incorporating both prior for smooth transition of

successive motion states and updates by the current sensory evidence (Khoei et al., 2017). Such computation can be implemented by reentrant feedforward connections, creating interference between an early and a late stimulus (Shimojo, 2014). Even if the mechanisms underlying the flash-lag effect remains speculative, both low and high-level processes could contribute (Hubbard, 2014).

3.2.2. Postdiction differs from retrocueing and retroperception

Backward masking is also a phenomenon where post-stimulus information influences perception. Two stimuli presented in close temporal proximity result in impaired visibility of the first stimulus (the target) due to the following stimulus (the mask). The temporal delay between target and mask leading to a maximized masking effect depends on the type of masking but is roughly around 100ms (Enns & Di Lollo, 2000).

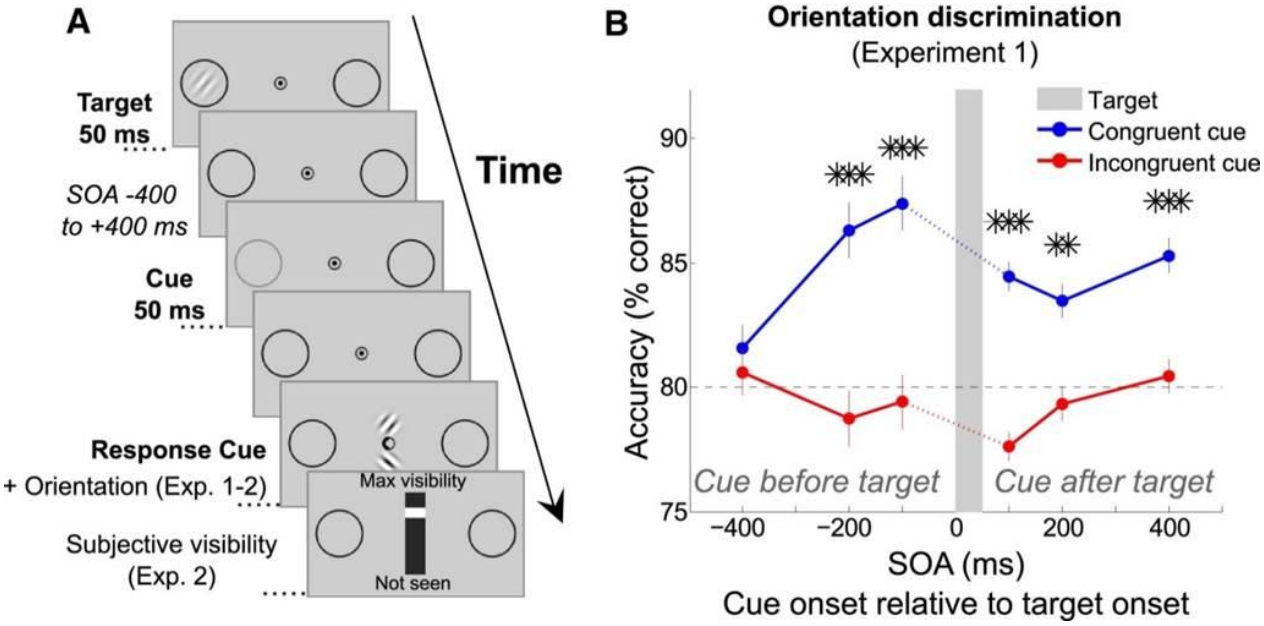


Fig. 3.5 Retroperception A. A cue indicating the target side (grey contour) was presented before or after the target (Gabor). Participants reported the direction of the tilted Gabor and rated their subjective visibility. **B.** Accuracy increased when the cue is congruently presented -200 ms before the target (shaded area), but also up to 400ms after (blue), compared to the incongruent cueing condition (red). This improvement in conscious perception due to postcueing is called retroperception. (Adapted from Sergent et al., 2013).

A second stimulus closely following a target does not necessarily impair the target perception but can have a beneficial influence, an effect called retro- or post-cueing. In both pre- and retro-cueing, alpha desynchronization was observed contralateral to the cued side, even if the ERP patterns differ, suggesting that the early evoked response integrates pre-stimulus influence only (Myers et al., 2015). Retrocueing was thought to selectively favor the transfer of sensory information into short-term memory. A recent study challenged this view and proposed that retrocueing (or “retroperception” in the

authors' terminology) can affect conscious perception (Sergent et al., 2013). Indeed, a cue spatially congruent with the target and presented up to 400ms after a target (Fig. 3.5A) triggered the conscious perception of the target that would have otherwise been missed (Fig. 3.5B). Retroperception has been proposed to be accounted for by selective attention which would retrospectively be attracted and consequently retrospectively enhance visual memory traces, prompting a new conscious percept (Sergent et al., 2011; Sergent et al., 2013, Thibault et al., 2016). An alternative explanation would be that retroperception is partially triggered by sensory-level interactions. A condition where both sides were cued showed the retroperception effect up to 200 ms (Sergent et al., 2013), which could be explained either by low-level interactions or by an attentional focus on multiple separate locations. To control for this, the visual properties of the cue were changed: the modification did not have any impact on retroperception, which discards the sensory-level hypothesis. Yet, a recent study challenges this view and claims that retroperception could be triggered by pure sensory-level integrations, without any redirection of spatial attention (Xia et al., 2016). The experimental design was almost identical to Sergent and colleagues' design (2013), except that the cue and the target were always presented at the same fixed position within a block, ensuring that full attention was directed to this position. Increased discrimination accuracy was observed for post-cue presented up to 150-200ms, a temporal range similar to the range observed in the bilateral cue condition in Sergent's study. However, subjective visibility did not correlate with objective accuracy, on the contrary to Sergent's study, which is not consistent with the notion of a retrospective triggering of conscious perceptual experience. Moreover, they also varied the visual properties of the cue, which resulted in decreased accuracy. The discrepancy with Sergent's study could originate from the more pronounced difference in cues (dot ring in Xia's study, noise-textured ring in Sergent's study). However, the sensitivity was not significantly different between the two types of cue. Conclusion about a possible sensory-level account for retroperception is hard to draw from these mixed findings and more investigation will be needed to rule on the matter.

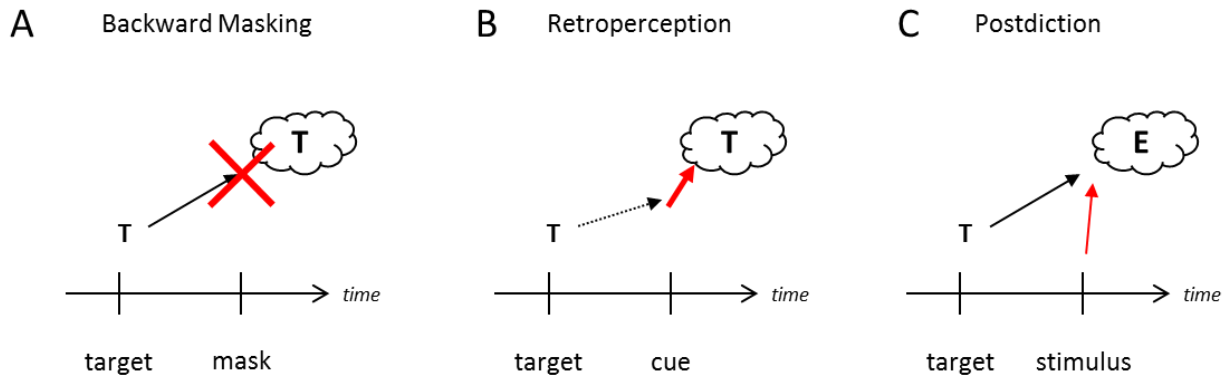


Fig. 3.6 Postdiction differs from backward masking and retroperception. Backward masking (A) and retroperception (B) modulate the conscious access of the target information, either impairing it (red cross) or enabling it (red arrow) in the case where the target information would have been under consciousness threshold (dashed arrow). On the contrary, the stimulus following a target in postdiction (C) does not influence its visibility but rather modify the content of the conscious representation (the feature T of the target is consciously perceived as feature E).

In postdiction, the conscious content of the target representation itself is modified, unlike backward masking or retroperception, where the target visibility is modulated by the following stimulus (Fig. 3.6). We postulate that postdiction can be formally singled out from retroperception and backward masking due to fundamental different mechanistic origins in particular with respect to the functional role of attention in these phenomena. Retroperception appears to be accounted for by a purely attention-driven mechanism triggering a new percept into consciousness (Sergent et al., 2013) – and maybe partially by low-level interactions (see previous paragraph) - whereas backward masking could be explained by independent low-level mechanisms such as contour-inhibitory processes based on lateral inhibitions (Breitmeyer & Ogmen, 2000). Postdiction can also be modulated by attention (Kilgard & Merzenich, 1995; Baldo et al., 2002) but its core mechanism is more likely pre-attentional (Nijhawan, 2002).

3.2.3. Postdiction as a general mechanism for perception

Postdictive phenomena were also reported at the level of memory for larger temporal scale ranging from minutes to months. Hindsight bias is for instance the tendency to judge past events as being more foreseeable that what they actually were, observed in judging historical, sport or trial-related events (Shimojo, 2014). In a controlled-lab experiment, the participants' choice was biased toward a stimulus-based judgement while they believed having made their choice *before* the presentation of the stimulus (Bear & Bloom, 2016). The similarity between perceptual and long-term postdiction could reflect shared functional processes homogenizing and revising knowledge/perception in order to achieve causal consistency, even if the underlying mechanisms are probably different. The idea that causality is postdictively reconstructed stems from the decorrelation between physical order and perceived causality (Choi & Scholl, 2006). Causality causes temporal compression (Haggard et al., 2002; Eagleman & Holcombe, 2002; Faro et al., 2013) and, even more strikingly, could drive order perception (Bechlivanidis & Lagnado, 2016) and event segmentation (Radvansky & Zacks, 2014). Postdiction is also at the core of the debate on free will launched by Benjamin Libet's experiments: the sense of agency could issue from a postdictive assessment on the consistency of a percept (Shimojo, 2014).

3.3. The Rabbit illusion

3.3.1. The Rabbit illusion as a proxy for postdiction

Saltation is a spatio-temporal illusion explained by postdiction, where a regular sequence of tactile stimuli is presented at two locations on the skin (Geldard & Sherrick, 1972). The two first stimuli are presented near to the wrist, while the last is presented near to the elbow (Fig. 3.3C). The second stimulus is perceived as mislocalized toward the position of the last stimulus, as if a little rabbit was hopping along the arm, hence the name of the Rabbit illusion. The presentation of the last stimulus influences the perception of the second one, making postdiction a prime explanation for saltation. Compared to the flash-lag effect, a minimal set of three stimuli is needed to elicit the illusion (Geldard, 1982), which allows for an easier neuroimaging investigation. The stimuli are of the same nature without any motion involved anymore, which discards de facto the latency-difference and the motion-extrapolation theories proposed for the flash-lag effect.

Moreover, the relative displacement observed between perceived locations depends only on the temporal delay between stimuli and is independent of attention (Fig. 3.7A), suggesting that postdiction is an automatic pre-attentional process. It is yet noteworthy that attention can modulate the absolute position of the stimuli, because it could explain the discrepancy between earlier results and recent replications (Kilgard & Merzenich, 1995). Indeed, Geldard reported a complete mislocalization of the second stimulus at the last stimulus' position (Geldard, 1977; Flach & Haggard, 2006), while following studies only found maximal mislocalization half-way between the two stimulated positions (Fig. 3.7B and C, Asai et al., 2012). The intensive training of Geldard's participants probably boosted the effect.

Last but not least, the Rabbit illusion was observed in tactile (Geldard, 1977; Geldard & Sherrick, 1983; Eimer et al., 2005; Flach et al., 2006; Trojan et al., 2009: Blankenburg et al., 2006), but also auditory (Hari et al., 1995; Shore et al., 1998: Phillips et al., 2001; Kidd et al., 2004; Ishigami & Phillips, 2008) and visual (Khuu et al., 2010; Khuu et al., 2011) modalities, strongly speaking in favor of a common central mechanism structuring perception. As in the flash-lag effect, the feature postdictively revised is the spatial location. Yet, postdiction was also assessed on changing color (Kamitani & Shimojo, 1999) during a colored flash-lag effect (Sheth et al., 2000). Overall, the Rabbit illusion appears as a convenient perceptual illusion to address postdiction.

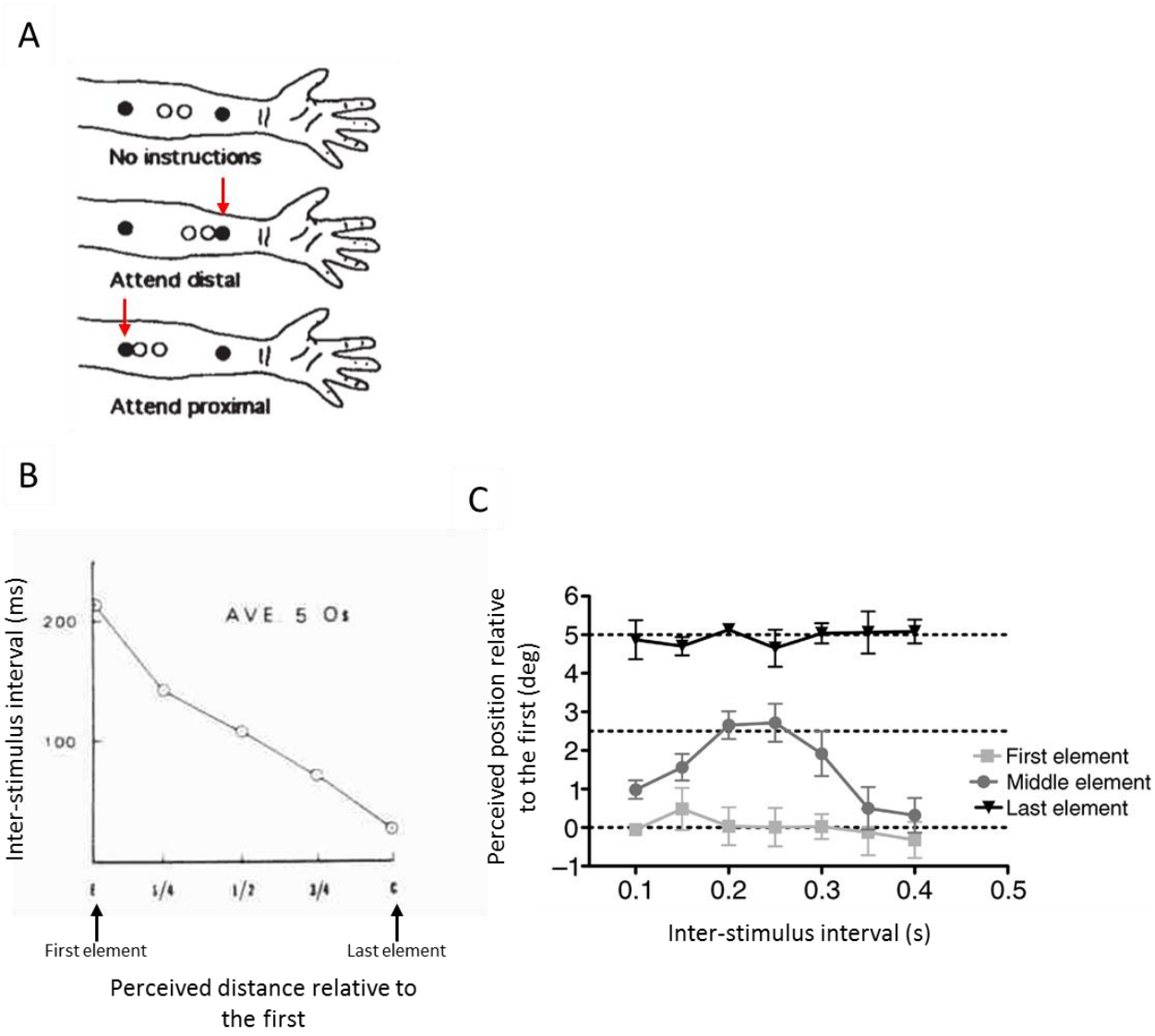


Fig. 3.7 Temporal profile of the Rabbit illusion. A. When a sequence of four stimuli are presented at two positions (two near to the elbow, two near to the wrist), attention (red arrow) modulates the perception of the intermediate stimuli. However, the distance between the intermediate stimuli depends only on the ISI and not on attention. **B.** In early reports, the mislocalization is total: the intermediate stimulus is perceived at the last stimulus' position when the ISI (inter-stimulus interval) is low. **C.** Such « full rabbit » were not reported afterwards, as the maximal mislocalization (medium gray) was found halfway between stimulated position (black and light gray) for intermediate ISIs. Intensive training and attention manipulation were advocated to explain the discrepancy. (Adapted from A: Kilgard & Merzenich, 1995; B: Geldard, 1977; C: Asai et al., 2012)

3.3.2. The Rabbit illusion depends on a "slow-speed" prior

One crucial feature of the mechanism underlying the Rabbit illusion is the involvement of a "slow-speed" prior knowledge, which was formalized in Bayesian terms (Goldreich & Tong, 2013, Khoei et al., 2017). This prior expressed the higher likelihood for smooth motions, probably learned from

natural statistics (Jones & Huang, 1982). Once combined with perceptual evidence, the posterior probability biases the percept in favor of a regular spatial distribution, resulting in the second stimulus apparently attracted by the last one. The prior would be based on two assumptions: (1) a single object is moving, and (2) the stimulus is temporally regular. Indeed, the violation of assumption 2, tested by introducing temporal irregularities in the sequence, weakens the illusion (Geldard & Sherrick, 1972). On the opposite, the effect was stronger for sequences with numerous stimuli compared to shorter ones (Geldard, 1977; Flach & Haggard, 2006), which strengthen both assumptions. Besides, when vertical and horizontal gratings are used, the intermediate stimulus is strikingly following either a linear or a curvilinear trajectory consistently with the grating orientation (Khuu et al., 2011; Fig. 3.8). The visual features of the gratings are thus integrating into an object representation, used as the basis to infer the most probable trajectory.

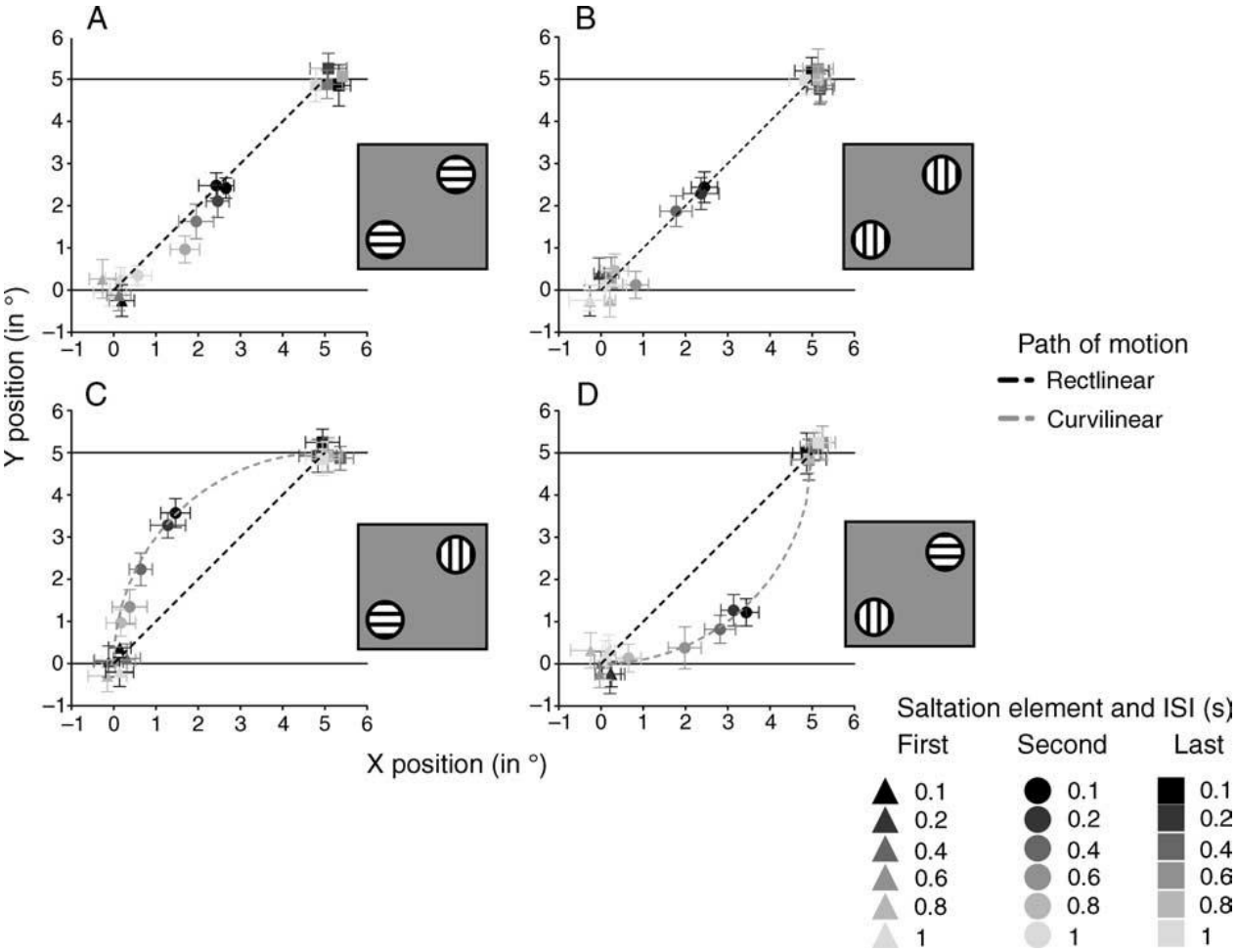


Fig. 3.8 Prior information involved in the rabbit illusion. Depending on the orientation of the stimuli gratings, the spatial mislocalization of the second stimulus will follow a linear (A and B) or a curvilinear (C and D) trajectory. The shade of gray represents different inter-stimulus intervals (ISI). (Adapted from Khuu et al., 2011).

If postdiction indeed relies on a prior probability, increasing the uncertainty of sensory encoding should strengthen the illusion as this will confer stronger weight to the prior. The flash-lag effect was for instance enhanced with blurred rather than sharp stimuli (Soga et al., 2009), larger eccentricity, faster target velocity, and less predictability on the presentation of the flash object (for a review, see Hubbard, 2014). Larger flash-lag effects were also reported when the flash occurs at a specific phase of pre-stimulus alpha oscillations inhibiting stimulus encoding (Chakravarthi & VanRullen, 2012). Altogether, these evidences are in line with the involvement of prior knowledge in postdiction, which is a key process differentiating postdiction from retroperception and backward masking. The functional relevance of postdiction could be to assure the consistency of perception, filling-in the blanks let by noisy sensory evidence, as the Rabbit was even felt even in areas lacking of receptors (Lockhead et al., 1980; Geldard et al., 1977; Miyazaki et al., 2010).

3.3.3. Neural correlates

One fMRI study investigated the neural correlate of the tactile Rabbit illusion (Fig. 3.9A). They showed that the retinotopic area corresponding to the perceived but not stimulated location was activated by the illusory sequence (Blankenburg et al., 2006; Fig. 3.9B). Higher-order areas as right premotor and prefrontal cortices were also activated by the illusion (Fig. 3.9C). Concerning the exploration of temporal dynamics, disruption of the activity in MT+ significantly reduced the flash-lag effect when TMS pulses are send 200ms after the flash onset (Maus et al., 2012).

The activation of high-level areas is consistent with the notion of prior in the illusion; however the underlying mechanisms remain unclear. Top-down feedback projections have been related to alpha/beta bands while feedforward information is carried by gamma oscillations (van Kerkoerke et al., 2014; Michalaeras et al., 2016; Sedley et al., 2016). Cross-frequency coupling enables the multiplexing of top-down and bottom up information (Fontolan et al., 2014). Moreover, cross-frequency coupling was modulated by learning ordering relationship, in monkeys and humans' auditory areas (Kikuchi et al., 2017). We hypothesized that an increased activity in low frequency bands and a potential change in cross-frequency coupling could translate the involvement of prior information in the illusion.

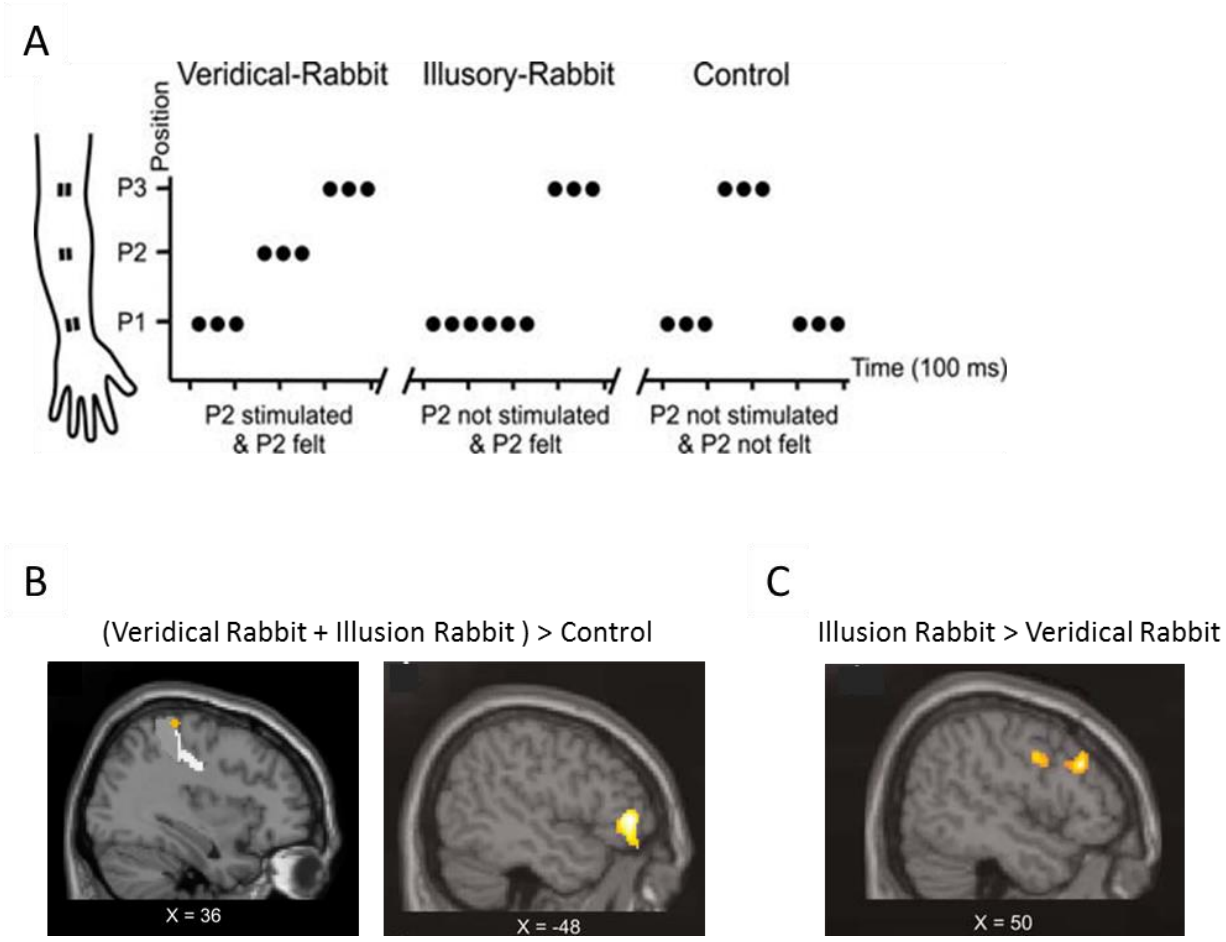


Fig. 3.9 fMRI activation during a tactile rabbit A. Three conditions were contrasted: veridical rabbit, illusory rabbit and control. **B.** Common activation for the veridical-rabbit and illusory-rabbit conditions are shown in orange (left image), projected onto the localizer contrast (P1 versus P2 and P3 in bright gray, P2 versus P1 and P3 in intermediate gray, P3 versus P1 and P2 in dark gray). Both Rabbit conditions activate the areas coding for P2. In the whole brain analysis (right image), the same contrast showed activation of the left inferior frontal gyrus. **C.** When the illusory rabbit is contrasted to the veridical rabbit, the right middle frontal gyrus and the precentral inferior frontal gyrus are activated. (Adapted from Blankenburg et al., 2006).

3.4. Summary and working hypotheses

Encoding windows entail that perception of a brief stimulus arises from the combined influence of sensory evidence temporally flanking the target. Pre-stimulus information such as predictions may constrain the perceptual inference made on subsequent stimuli. Predictions were hypothesized to be carried by low-frequency oscillations. Postdiction consists in the influence of a late stimulus on the conscious representation of a stimulus that was presented earlier. The perceptual change in the representation of the event may be mediated by the influence of prior knowledge learned from

ecological statistics. One hypothesis is that top-down expectations, supported by low-frequency oscillations, can be related to prior information, while feedforward sensory evidences could be coded by gamma activity. The Rabbit illusion appears as a convenient tool to explore postdiction as it involves a simple sequence of three identical stimuli.

Prediction and postdiction builds perception

3.5. Goal, hypothesis and summary of results

Ongoing neural activity is now widely recognized for impacting the processing of upcoming stimulus. More surprisingly, late stimulus can also influence early stimulus. Postdiction precisely consists in the possibility that a later stimulus influences the conscious representation of a stimulus that occurred before it. The representational updating was thought to arise from top-down expectations based on prior knowledge. High-order areas such as right premotor and prefrontal cortices were indeed related to postdiction. However, top-down expectations from higher-order areas typically reflect predictive mechanisms modulating low-level sensory areas. The interplay between prediction and postdiction can only be unraveled with time-resolved imagery, which was left unexplored up to now.

To investigate the underlying mechanisms of postdiction, we used a visual version of the Rabbit illusion (Geldard, 1977). An isochronous sequence of three flashes was presented in the periphery of the visual field, with the two first flashes presented at the same location, and the last one a little further away. The second stimulus was perceived in-between the physically stimulated locations ~50% of the time following the calibration of the inter-stimulus interval. Hence, the same visual sequence was either correctly perceived or illusory, enabling us to compare different percepts elicited by the same stimulation. In a control condition, the intermediate stimulus was presented at an intermediate position yielding the same perceptual report as the rabbit illusion. To balance the participant's attentional focus, two conditions were added that changed the direction of the visual sequence (inward or outward).

Overall, perceiving the rabbit illusion was associated with increased amplitude of the evoked responses following the presentation of the full visual sequence in the left premotor area and the right parieto-frontal regions along with a differential activation in the right retrosplenial/primary visual regions. The analysis of oscillatory activity preceding the presentation of visual sequences showed that an increased alpha power in visual regions was correlated with the illusory perception of the sequence, but only in the outward condition. Both pre- and post-stimulus activity seems to influence the perception of the second stimulus embedded in the sequence. More investigation is needed to determine the oscillatory mechanisms underlying postdiction, and to assess whether prediction plays a role in the rabbit illusion. Analyses are therefore ongoing at the time of writing.

3.6. Introduction

Sensory information unfolds over time, along with the neural and cognitive operations that process the sequence of events. One hypothesis is that the temporal structure of sensory events unfolding over time is itself an information that the brain incorporates in its sensory analysis to make sense of its environment (Zacks & Tversky, 2001). Under this working hypothesis, the perceptual outcome is delivered at the end of what has been recognized by the system as a sequence of temporally related stimuli. One question is how the brain determines the parsing size of informational chunking and this has been largely addressed for speech (Giraud & Poeppel, 2012). Another question is that under this scheme, the perception of a stimulus in a sequence may be affected by the latest stimulus in the sequence. The perception of earlier stimuli will therefore seem to be causally driven by latest stimuli. This phenomenon, called postdiction, may be related, for instance, to situations where a mask impairs the visibility of a preceding target (backward masking) or rather improves it (Sergent et al., 2013). Postdiction is thought to be the source of perceived apparent motion or the rabbit illusion (Bachmann, 2013; Shimojo, 2014). In the rabbit illusion, earlier stimuli within a sequence are mislocalized due to the temporal regularity of the full sequence (Geldard & Sherrick, 1972). For instance, if two tactile stimuli presented on the wrist are followed by one stimulus presented near to the elbow, the second stimulus will be perceived at an intermediary position between the first and last stimulus in the sequence, as if it was attracted towards the last stimulus in the sequence. This illusion was described in the tactile (Geldard, 1977; Geldard & Sherrick, 1983; Eimer et al., 2005; Flach et al., 2006; Trojan et al., 2009: Blankenburg et al., 2006), but also in the auditory (Hari et al., 1995; Shore et al., 1998: Phillips et al., 2001; Kidd et al., 2004; Ishigami & Phillips, 2008), the visual (Khuu et al., 2010; Khuu et al., 2011) and across sensory modalities (Kamitani & Shimojo, 2001), suggesting a common central mechanism temporally structuring perception.

Despite the importance of the temporal dynamics underlying the postdiction phenomenon, surprisingly little is known with regards to the neural implementation of postdiction and very few neuroimaging studies have tackled the question (Blankenbourg et al., 2006; Maus et al., 2012). For instance, postdiction has been proposed to explain a visual illusion called the flash-lag effect, in which a moving object is perceived ahead of a flash, although both are presented simultaneously (Nijhawan, 1994). Postdiction would intervene in the flash-lag effect because a change in the *post-flash* trajectory of the moving object, while keeping an identical pre-flash trajectory, was influencing the final percept, ruling out pure predictive accounts (Eagleman & Sejnowski, 2000). TMS applied to motion area MT+ has been shown to significantly reduce the perceived lag when TMS pulses were sent 200 ms after the flash onset (Maus et al., 2012), indicating that post-flash information were integrated in the perceptual representation.

In this study, we wished to investigate the “when” of postdiction with combined magneto- and electro-encephalography (M/EEG) recordings. However, the use of a moving stimulus as in the flash-lag effect prevents a temporally precise tracking of information in brain activity. On the contrary, the visual

rabbit illusion reduces the flash-lag effect to the more simple expression of a 3-flash sequence, which shrinks the complexity of the phenomenon while preserving the core operations. The discretization of the sequence allows, for instance, disentangling evoked brain signals responses. A sequence of three flashes was presented on the periphery of the right visual field. The stimulus onset asynchrony (SOA) was chosen so that the illusion was seen half of the time (test trials). This provided us with two post-hoc conditions deriving from an individual's perceptual outcome given the sequence: illusory (ILL) and veridical (VER) trials in which participants mislocalized or correctly reported the localization of the second stimulus in the sequence, respectively (Fig. 3.11A). Hence, ILL trials with VER trials provided a clear contrast of the sensory and decisional processes involved in the illusory perception given the same physical stimulation. Additionally, a control condition (CTRL) was introduced in which the second stimulus was physically presented at an intermediary position between the first and last flashes in the sequence. The CTR trials thus consisted in a similar perceptual outcome as the ILL trials, but with a different visual sequence.

More specifically, we build on the working hypothesis that the reconstruction of the visual sequence in the rabbit illusion is effectively postdictive, meaning that "a stimulus presented later seems causally to affect the percept of another stimulus presented earlier", following the operational definition of postdiction proposed by Shimojo (2014). Two predictions derive from this working hypothesis: (1) the last stimulus may trigger the activation of neurons that are retinotopically coding for the filled-in position, and (2) higher brain areas may be involved in the reallocation of spatial loci for each stimulus after the sequence presentation. In the tactile modality, the somatotopic area corresponding to the perceived but not stimulated location was indeed activated by the rabbit illusion sequence (Blankenburg et al., 2006). Higher-order areas as right premotor and prefrontal cortices were also activated by the illusion. Our two predictions would extend Blankenburg et al.'s proposals (2006) to the visual modality and further explore the temporality of the process. M/EEG source reconstruction was used to estimate the neuroanatomical loci of significant differences both for evoked and oscillatory brain responses.

From a computational viewpoint, the rabbit illusion can be scrutinized under a Bayesian framework. The perceived position of a stimulus combines both the likelihood estimating spatial uncertainty and a prior probability on spatial location given the frequency of presentation. This prior expresses the strong expectation toward stimuli moving at slow speed, probably learned from natural statistics (Jones & Huang, 1982; Goldreich & Tong, 2013, Khoei et al., 2017). The resulting posterior probability hence favors the perception of a regular spatial distribution. Based on this account, increasing the uncertainty of sensory encoding should result in an illusory perception, as the system is relying more on the prior probability. The flash-lag effect was for instance enhanced with increased sensory uncertainty (Soga et al., 2009). A possible marker indexing the level of excitability of a neural network is the spontaneous alpha oscillation (Romei et al., 2008; Iemi et al., 2016). Increased alpha power was proposed as inhibiting information processing (Jensen & Mazaheri, 2010). Hence, lower pre-stimulus alpha power could predict a veridical perception whereas an overall higher alpha power would predict an illusory perception.

Not only the power but also the phase of spontaneous alpha fluctuations modulates the excitability of neural populations (Klimesch et al., 2007; Haegens et al., 2011). If a stimulus arrives during a low excitability period (e.g. the peak of the alpha oscillation), the reliability in sensory encoding will decrease. The sensory uncertainty will thus be maximal, leading to lower stimulus detection (Mathewson et al., 2009; Busch et al., 2009; Dugué et al., 2011). The phase of 7Hz pre-stimulus activity predicted the strength of the flash-lag effect, suggesting that perceptual evidences are preferentially encoded during certain temporal windows (Chakravarthi & VanRullen, 2012). We thus tested the impact of the reliability of the first stimulus encoding on the processing of subsequent stimuli by contrasting alpha/theta phase between illusory and veridical trials (Fig. 3.10A). An alternative hypothesis considers that sensory inputs occurring within the same oscillatory cycles will be integrated together (Varela, 1981; Pöppel, 1991; Wutz & Melcher, 2014). In the rabbit illusion, if the second stimulus lied within the same encoding window than the last stimulus, the positional information of the latter will influence the perceived position of the former, leading to the illusion. However, if the second stimulus lied within the same encoding window than the first stimulus, it is more likely to be perceived at its veridical position (Fig. 3.10B). We expected that the phase of at the entrained frequency (around 4Hz) at the first stimulus onset can predict the illusion.

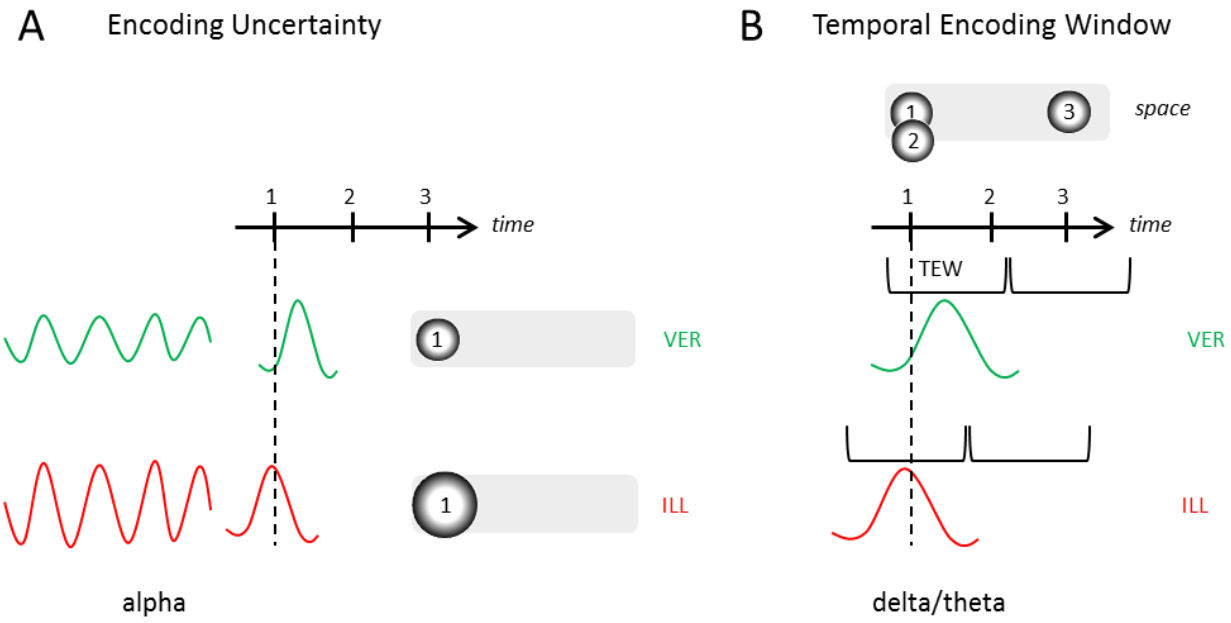


Fig. 3.10 Pre-stimulus phase and power hypothesis. A. In a Bayesian framework, increasing the sensory uncertainty on spatial position (depicted by a larger black and white circle) favors the illusory mislocalization of a flash embedded within a regular sequence. Both the power and phase of the alpha oscillation influence the proper encoding of incoming stimuli. We hypothesized that decreased alpha power, a marker of selective attention , predicts veridical perception (green) while increased alpha power predicts the illusion (red, left panel). Specific alpha phase enhance neural excitability and visual detectability. We hypothesized that the alpha phase corresponding to the onset of the first stimulus affects the spatial ambiguity on location of the first stimulus, hence biasing the perception of the whole sequence (right panel). **B.** An alternative hypothesis considers an oscillatory cycle as a temporal window of encoding (TEW). If the second stimulus lied within the same encoding window than the last stimulus, the positional information of the latter will influence the perceived position of the former, leading to the illusion (red). However, if the second stimulus lied within the

same encoding window than the first stimulus, it is more likely to be perceived at its veridical position (green). A difference of phase in the delta/theta frequency range is expected when contrasting the two percepts.

3.7. Materials and Methods

3.7.1. Participants

19 right-handed participants (8 males; age = 28 years, SD = 6) took part in the study. All had normal or corrected-to-normal vision, were not under any medical treatment, had no neurologic history, and were naive as to the purpose of the study. Each participant provided a written informed consent in accordance with the Declaration of Helsinki (2008) and the Ethics Committee on Human Research at NeuroSpin (Gif-sur-Yvette, France). Five participants were excluded from the study: the digitalization procedure prior to M/EEG recording was not carried out properly for one participant; the data for one participant was too noisy, and three participants did not reach the 60% performance criterion in the control task for one or two conditions. Hence, a total of 14 participants (6 males, age = 29 years, SD = 6) were considered in the final M/EEG analyses.

3.7.2. Experimental design

The stimuli consisted of three M-scaled 2D Gaussian blobs presented as a spatiotemporal sequence along the right horizontal meridian at 15.1°, 18.5° and 21.8° of visual angle (the diameter being respectively 2.5°, 3° and 3.5°). The second stimulus was presented at the same position than the first stimulus (test trials), while in the control condition, it was presented at an intermediate position between the first and last stimuli (Fig. 3.11A). Each stimulus was presented during 50 ms. The inter-stimulus interval (ISI) between the flashes was previously determined in a pilot study (n = 5) to cause the perception of the illusion in 50% of the trials. Knowing that the rabbit illusion can be influenced by attentional strategies (Kilgard & Merzenich, 1995), the sequence of flashes was presented centrifugally or centripetally relative to the fixation dot (Fig. 3.11B). Thus, participants could not predict where the next sequence started, ensuring that they payed equally attention to the three spatial locations. The SOA was 188 ms for the centripetal sequence of stimuli (inward condition) and 272 ms for the centrifugal sequence (outward condition). The reasons of this temporal asymmetry remain unclear but it mirrors previous difference in illusory reports observed for the flash-lag effect between centrifugal and centripetal conditions (Kanai et al., 2004). Each test condition was repeated 300 times and each control condition 150 times, leading to 900 trials in total. The experiment was divided into 5 blocks, each one containing 60 repetitions of each test condition and 30 repetitions of each control condition presented in a randomized order. Each block lasted around 8 min, and participants were encouraged to rest 1-2 min during the breaks. They were told that the locations of the first and last flashes were always fixed. They performed a 2-alternative forced choice by judging whether the second flash was in an intermediate position between the first and last ones. To help them recognizing the second flash, it was made brighter

(contrast 1, compared to 0.4 for the first and last flashes). The next trial only began after a random waiting period comprised between 800 and 1300 ms triggered by the response button press.

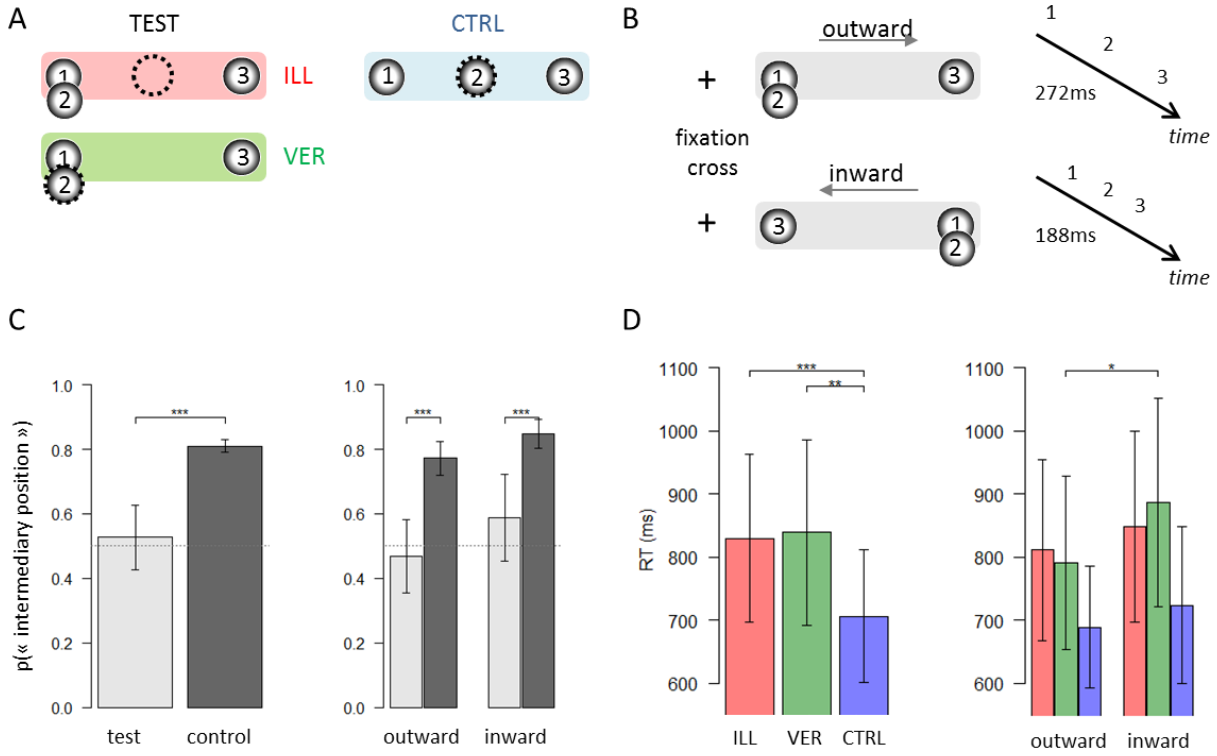


Fig. 3.11 Experimental design and behavioral results. A. In test trials, the two first flashes of the sequence were presented at the same spatial location. The second flash was perceived correctly (VER, in green), or at an intermediary position (ILL, in red). In control trials, the second stimulus was presented and perceived at an intermediary position (CTRL, in blue). Numbers in the circles indicate the temporal order of each flash within the sequence; dashed circle indicates the perceived position of the second flash. **B.** All stimuli were presented on the right horizontal meridian. Two directions for the sequence were possible, to ensure a balanced focus of spatial attention (outward or inward). The SOA was equal to 272ms in outward condition and 188ms in inward condition. **C.** The left panel shows that participants perceived the illusion in 53% (SEM=5) of test trials (light grey) and they perceived correctly 81% (SEM=0.9) of the control trials (dark grey). The right panel shows the same data split following on direction conditions. The dotted line corresponds to 50% of "intermediary position" response. **D.** In the left panel, reaction time was faster in CTRL trials (blue) than test trials (ILL, red, and VER, green). The right panel shows that reaction time is faster in outward than inward condition. *** p < .001, ** p< 0.01, * p < .05

3.7.3. Psychophysics analysis

Statistical analyses were conducted with R suite (R core team, 2013). A two-way repeated-measure ANOVA was conducted on response probability with the trial type (test, control) and condition (inward, outward) factors. A two-way repeated-measure ANOVA was conducted on reaction time with perceptual outcome (ILL, VER, CTRL) and condition (inward, outward) factors. Reaction times were computed from the onset of the last stimulus.

3.7.4. Combined M/EEG and anatomical MRI data acquisition

Brain activity was simultaneously recorded with magneto- and electro-encephalography (M/EEG) in a magnetically shielded room. A whole-head Elekta Neuromag Vector View 306 MEG system (Neuromag Elekta LTD, Helsinki) in upright position and equipped with 102 triple sensor elements (one magnometer and two orthogonal planar gradiometers per sensor location) and an Elekta EEG cap with 60 sensors were used. MEG recordings were sampled at 1 kHz and high-pass filtered at 0.3Hz. Horizontal, vertical and horizontal electro-oculogram (EOG) and electrocardiogram (ECG) were recorded during the session. The reference electrode for EEG was placed on the top of the participant's nose. Participants' head position was measured before each block with four head position coils (HPI) placed over the frontal and mastoid areas. Anatomical landmarks were used to locate the fiducial points (nasion and the left and right pre-auricular areas) and help for the coregistration between electrophysiological and MRI data. A 2 min MEG signal without participant in the room was recorded the same day for the computation of the noise matrix. After the end of the M/EEG session, the participant underwent an anatomical MRI using a 3-T Siemens Trio MRI scanner to provide high-resolution structural brain image. The classical MPRAGE sequence (voxel size: 1.0 x 1.0 x 1.1 mm; acquisition time: 466s; repetition time TR = 2300ms; and echo time TE= 2.98 ms) was used as well as the multi-echo FLASH sequence with flip angle at 5° and 30°, which ensured a better separation for skin and skull boundaries during the creation of the head model.

3.7.5. M/EEG data preprocessing

Signal space separation was applied to decrease the impact of external noise on recorded brain signals (Taulu et al., 2003). Signal space separation correction, head movement compensation, and bad channel interpolation were done using MaxFilter Software (Elekta Neuromag) after visual inspection of the raw data. Volumetric segmentation of participants' anatomical MRI and cortical surface reconstruction was performed with the FreeSurfer software (http://surfer.nmr.mgh.harvard.edu/). These procedures were used for group analysis with the MNE suite software (Gramfort et al., 2014). The co-registration of the MEG data with the individual's structural MRI was carried out by realigning the digitized fiducial points with MRI slices. Using *mne_analyze* within the MNE suite, digitized fiducial points were aligned manually with the multimodal markers on the automatically extracted scalp of the participant. To insure reliable coregistration, an iterative refinement procedure was used to realign all digitized points with the individual's scalp. Finally, ocular and cardiac artefacts were corrected by rejecting ICA components, computed separately for EEG and MEG data, which correlates with ECG/EOG events. Epochs were rejected if their amplitude exceeded a certain threshold (gradiometers: 4000 e^{-13} T/m, magnetometers: 4 e^{-12} T, EEG: 250 e^{-6} V). Bad EEG channels were visually detected and then interpolated.

3.7.6. M/EEG analysis

Trial selection

M/EEG data analysis was conducted with MNE-Python (Gramfort et al., 2013, 2014). Trials whose reaction time was faster than 200ms or slower than 4s were rejected (about 2% of rejection; RT from the offset of the second stimulus in the sequence). Trials were split following the perceptual outcome: in test trials, participants could perceive the illusion (ILL) or the correct localization (VER); in control trials, only correctly perceived trials were considered and thereafter labelled CTRL trials (81 +/- 0.9 % of trials).

Epoching

M/EEG data were low-pass filtered at 35 Hz and down sampled to 250 Hz. Epochs were locked on the onset of the first stimulus in the sequence from -300 to + 1200 ms. Baseline-correction was computed on the prestimulus period (-300 to 0 ms). Trials were sorted as a function of participants' response and the number of epochs was equalized across conditions using automatized procedure in MNE-Python which insured that the selection of trials across conditions was as close in time during the acquisition as possible (cf. the function mne.epochs.equalize_epoch_counts on http://martinos.org/mne). In the inward condition, 76 trials (SD = 30) and 107 trials (SD = 59) were preserved on average when contrasting ILL *vs*. VER and ILL *vs*. CTRL, respectively. In the outward condition, 91 trials (SD = 37) and 105 trials (SD = 35) were used when contrasting ILL *vs.* VER and ILL *vs.* CTRL, respectively.

Source reconstruction

Individual forward solutions for all source locations located on the cortical sheet were computed using a 3-layers boundary element model (BEM) constrained by the individual's anatomical MRI. Cortical surfaces extracted with FreeSurfer were sub-sampled to 10,242 equally spaced sources on each hemisphere (3.1 mm between sources). The noise covariance matrix for each individual was estimated from the raw empty room MEG recordings preceding the individual's MEG acquisition. The forward solution, the noise covariance and source covariance matrices were used to calculate the dynamic Statistical Parametric Mapping (dSPM) estimates (Dale *et al.*, 2000). The inverse computation was done using a loose orientation constraint (loose = 0.4, depth = 0.8) on the radial component of the signal. Individuals' current source estimates were registered on the Freesurfer average brain for surface based analysis and visualization.

Labels selection in source space

Labels were selected from PALS_B12_Brodmann (BA) and PALS_B12_Visuotopic parcellations on the basis of the activity of combined ILL, VER and CTRL trials (Fig. 3.12). The visual activity was captured by V1, VP, middle temporal (MT) but also by BA30 as the visual response was medially spread. After S2 presentation, temporal areas as the supramarginal (BA40) and the angular gyrus (BA39) were activated. At the end of the sequence, orbitofrontal areas (BA10), posterior cingulate (BA31) and motor cortex (BA4) were also activated. In addition, BA8 and BA10 were also selected to replicate Blankenburg et al.

(2006) analysis which found different activations in the inferior frontal cortex and middle frontal gyrus. A total of 22 labels (11*2 hemispheres) were used for the analysis.

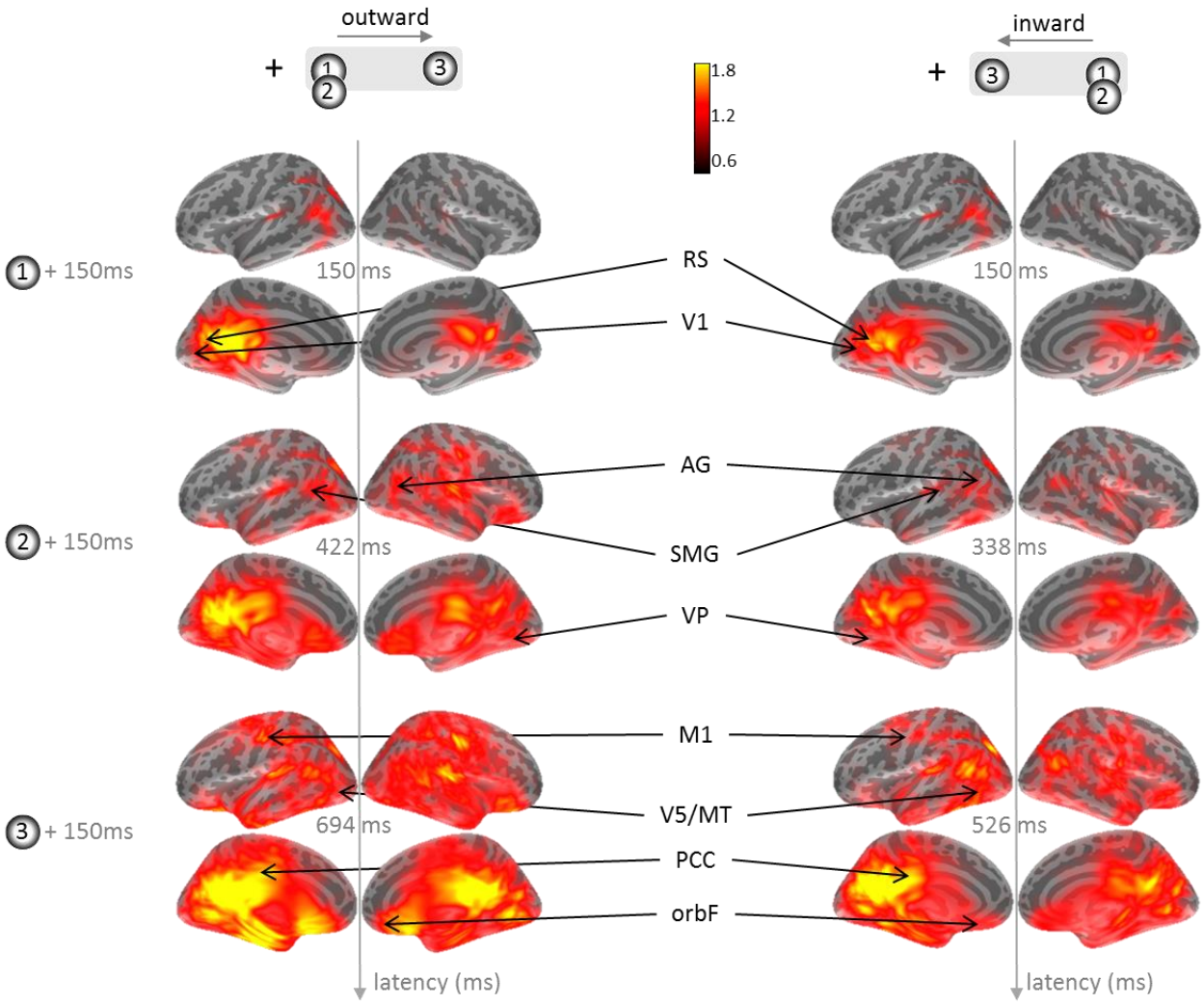


Fig. 3.12. Label selection. Source-reconstructed activity was averaged across control and test trials and plotted for each direction condition (inward on the right, outward on the left). Abbreviations: BA: Brodmann area, RS: retrosplenial (BA 30), V1: primary visual area, AG: angular gyrus (BA 39), SMG: supramarginal gyrus (BA 40), VP: ventral occipital, M1: primary motor area (BA 4), V5/MT: medial temporal, PCC: posterior cingulate cortex (BA 31), orbF: orbitofrontal (BA 10)

Statistical analysis of evoked responses

We contrasted ILL and VER trials to reveal differences in perceptual processes, and ILL and CTRL trials to investigate postdictive processes. Permutation paired t-test with spatio-temporal clustering (cluster threshold = 0.05) was realized on sensors data for each type of channels, from S2 onset up to 800 ms. In source space, permutation paired t-test with temporal clustering (cluster threshold = 0.05) was performed on each selected label. Only clusters whose p-value was above .05 were reported.

Phase analysis

Two alternative hypotheses predict a phase difference between illusion and veridical perception. First, the alpha/theta phase at the onset of the first stimulus could determine the quality of encoding of the first stimulus, hence the reliability of the whole sequence encoding. Second, low-oscillation cycles could determine temporal window of encoding. The instantaneous alpha phase was computed for each epoch at the onset of the first stimulus. Morlet wavelets were centered on 10Hz for the alpha band, 7Hz for theta band (as in Chakravarthi & VanRullen, 2012) and 4Hz, with respectively a number of cycles of 3.3, and 2.5. Phase analyses were carried on the sensor showing maximal pre-stimulus power from -600 to -100 ms for test trials. Inward and outward conditions were combined for this analysis. For each participant, all phases were normalized relative to the mean phase of ILL. At the group level, a Rayleigh test assessed whether each condition showed a preferential phase. To evaluate whether the two distributions were significantly different, a surrogates distribution was created by randomly shuffling the epochs' labels (N = 1000). The normalization procedure was applied for each participant and the difference of phase was computed between the two random distributions at the group level. A 95% confidence interval was derived from this distribution. If the averaged phase difference between ILL and VER was outside the confidence interval, it was considered as significant at p = .05.

Time-frequency analysis

Epochs were defined from -800 to +1200 ms post-stimulus, downsampled to 333Hz and low-passed filtered at 160Hz. Time-frequency analysis was applied on the whole epochs and on each sensor. Power was calculated using Morlet wavelets for the alpha band (7-13Hz) with 6.7 cycles. As the time-frequency hypotheses focused on alpha pre-stimulus activity, no baseline was applied for the computation of oscillatory activity (Jensen & Mazaheri, 2010; Busch et al., 2009; Hanslmayr et al., 2007). Alpha power was computed for combined in- and out-ward conditions but also for each condition separately. Spatio-temporal clustering based on permutation t-test was performed from - 600 to -100 ms post-stimulus onset for each channel type. Only significant clusters with p-value under .05 were reported. Alpha power was also computed in source space for each time point with Morlet wavelets (7-13Hz, 6.7 cycles).

3.8. Results

3.8.1. Behavioral results

On average across the inward and outward conditions, participants mislocalized the second stimulus in 53 +/- 5 % of the test trials and correctly perceived the location of the second stimulus in 81 +/- 0.9 % of the control trials (Fig. 3.11C). A two-way repeated-measure ANOVA was conducted on the probability of "intermediary position" responses with condition (test, control) and direction (inward, outward) as factors. A main effect of condition ($F(1,13) = 32.52$; $p = 7.10^{-5}$, $\eta^2_p = 1.562$) was found: participants perceived the second stimulus in intermediary position significantly more often in control

than in test trials, which was expected by experimental design. No effect of direction ($F(1,13) = 2.80$, $p = .118$, $\eta^2_p = .431$) nor interaction between condition and direction were found ($F(1,26) = 1.56$, $p = .233$).

Reaction times (RT) were also investigated with the working hypothesis that postdiction may implicate additional processing steps in ILL as compared to CTRL, which would be related to the temporal reorganization of events for the perception of the visual sequence. A two-way repeated-measure ANOVA was thus conducted on RTs, with direction (inward, outward) and perceptual outcome (ILL, VER, CTRL) as factors (Fig. 3.11D). A main effect of perceptual outcome was found ($F(2,13) = 5.703$, $p = .009$), but no significant effect of direction ($F(1,13) = 2.377$, $p = .147$) nor interaction between factors ($F(2,26) = 0.926$, $p = .409$) was found. A post hoc Bonferroni-corrected paired t-test showed that RT in CTRL was on average 123 ms faster than in ILL ($p = 10^{-4}$, $\eta^2_p = .512$) and 133 ms faster than VER ($p = .002$, $\eta^2_p = .534$). As expected, RT was slower in ILL compared to CTRL, which could be due to additional processing steps related to postdiction. However, RT was also slower in VER compared to CTRL, while in these two types of trials, the participant's response was correct. One explanation could be that participants actually perceived a small mislocalization even in VER trials, but the extent of the perceived mislocalization was not large enough to exceed a certain decisional threshold, resulting in a correct report. This implies that the temporal reorganization of events did occur in all test trials (hence the same difference of RTs in VER than ILL compared to CTRL) but in a lower extent than in ILL trials (hence the difference of report). An alternative post-hoc hypothesis could be that slower RT in VER as compared to CTRL trials may be due to additional perceptual or decisional processes dealing with the larger uncertainty of the sensory outputs in test trials.

3.8.2. Neural markers of postdiction

If the sequence was postdictively perceived, both visual and higher-order areas may be activated after the last stimulus presentation. Spatio-temporal clustering with permutation t-test was performed in sensor space on the two contrasts of interest, ILL vs. VER and ILL vs. CTRL. In central EEG channels, the neural activity was significantly different between ILL and CTRL, after the visual sequence (380-600ms, $p = .002$) but only in the inward condition (Fig. 3.14A). No other significant cluster was found, maybe because of the fact that MEG data were not realigned between participants. To overcome this problem, we performed temporal clustering with permutation t-test in source space within labels selected independently from the contrasts of interest (see *Labels selection in source space* in the methods section for more details). Common differences between in- and out-ward conditions were most likely involved in the common postdictive mechanism at stake. We nevertheless reported the differences specific to one condition.

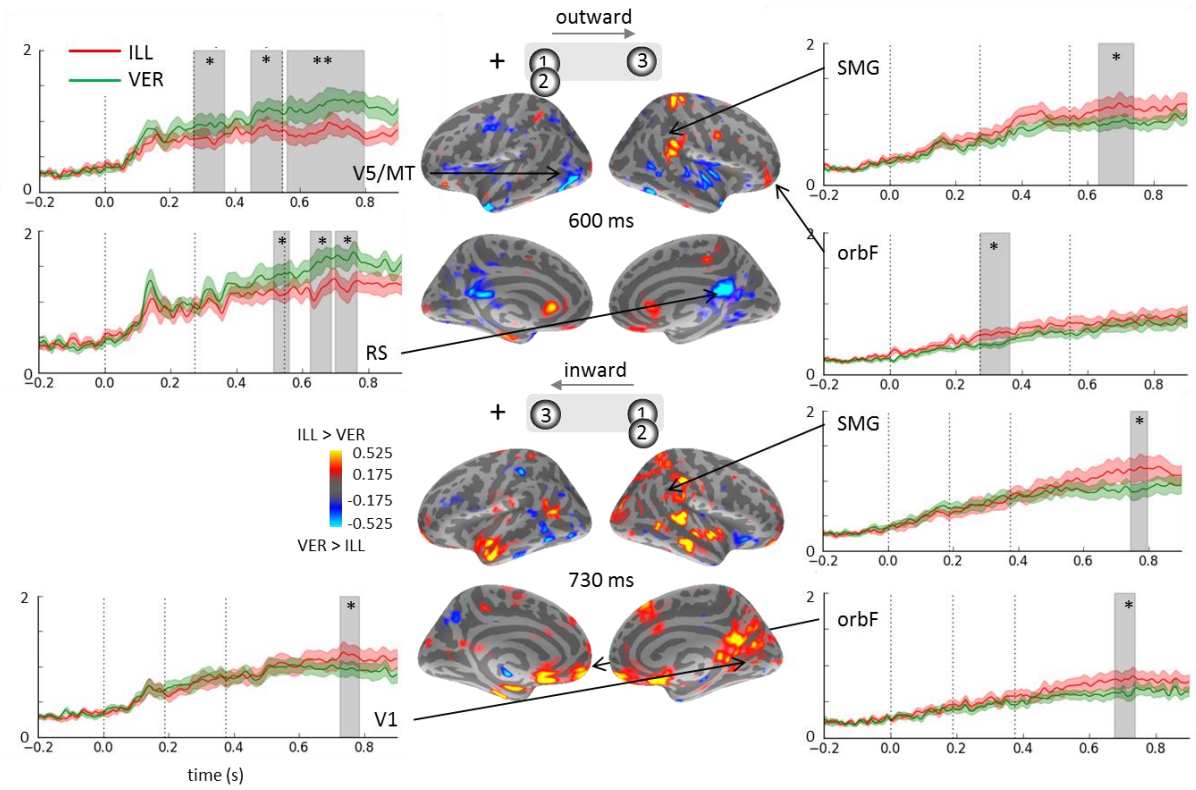


Fig 3.13 Contrasting different perceptions for the same stimuli. The contrast ILL versus VER was expected to inform on perceptual and decisional processes related to the reorganization of the sequence. We found common differences between out- (top panel) and in-ward (bottom panel) conditions in sensory areas (right visual/retrosplenial cortices) and in higher-order areas (supramarginal and orbitofrontal/inferior frontal gyrus). Specific difference was found in the outward condition in MT. p-values: * < .05, ** < .01
Abbreviations: BA: Brodmann area, RS: retrosplenial (BA 30), V1: primary visual area, SMG: supramarginal gyrus (BA 40), V5/MT: medial temporal, orbF: orbitofrontal (BA 10).

The comparison between ILL and VER may inform on differences in sensory and decisional processes. Common differences to direction conditions were found in right fronto-parietal areas (Fig. 3.13): the activity increased when the illusion was perceived in orbitofrontal areas (outward: rh, p = .015, 272-364 ms; inward: lh, p = .015, 676-736 ms) and in the right supramarginal (outward: p = .023, 632-734 ms; inward: p = .043, 744-796 ms). These high-order areas were found after the end of the sequence presentation, suggesting that they could be involved in postdictive mechanisms. Specifically to the outward condition, motion area MT was more activated in VER than ILL after S2 presentation (p = .025, 272-368 ms; p = .025, 448-544 ms; p = .009, 560-796 ms). A common area in the right hemisphere, overlapping between the retrosplenial cortex and V1, showed an opposite pattern between outward and inward condition: in outward condition, the activity was increased in VER (BA30-rh, p = .047, 512-560 ms, p = .017, 624-688ms, p = .016, 700-764 ms), while in the inward condition, it increased in ILL (V1-rh, p = .027, 724-784 ms).

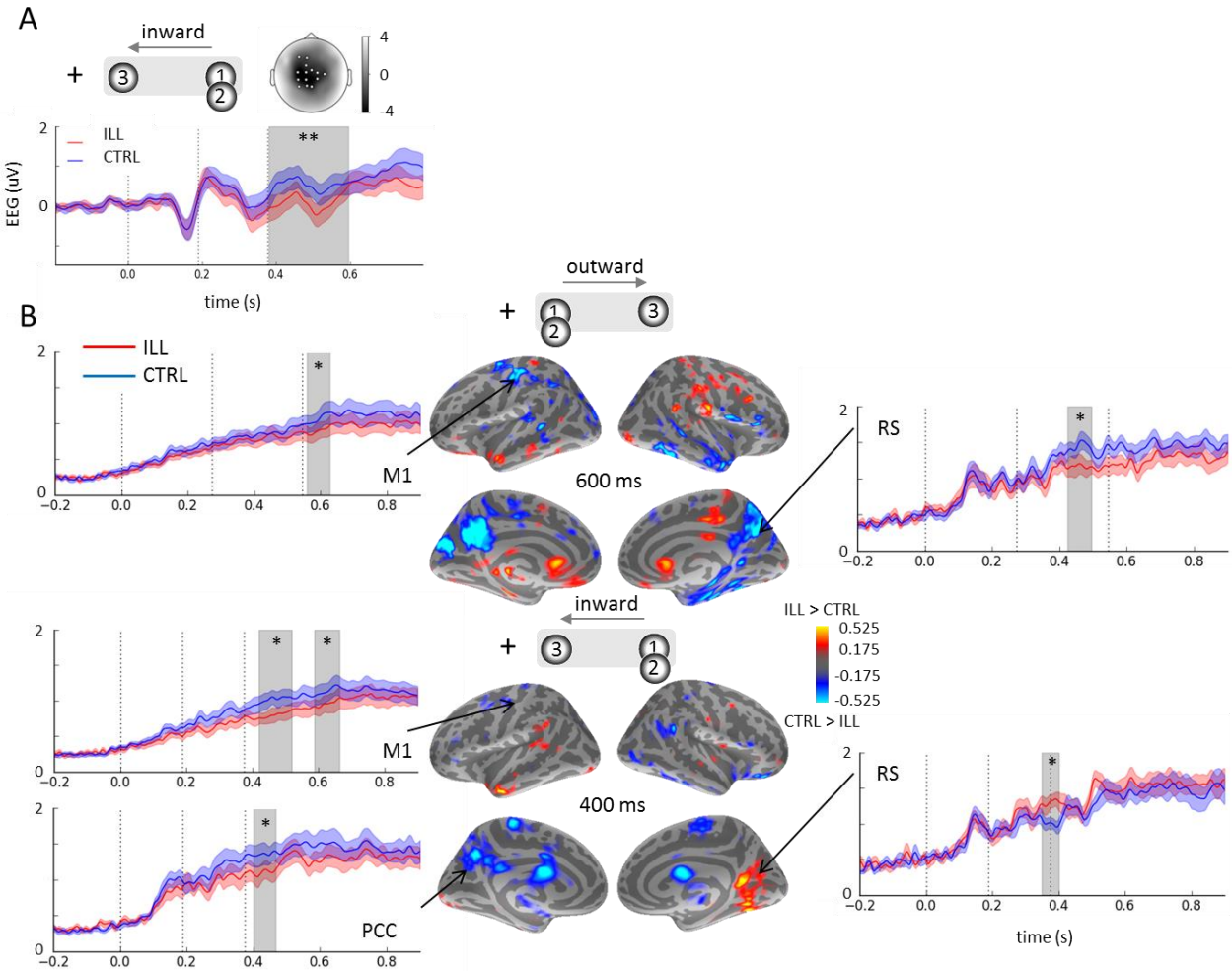


Fig 3.14 Contrasting the same perception for different stimuli. The contrast ILL versus CTRL was expected to inform on the timing of postdiction. **A.** Sensors space analysis indicates significant difference in central sensors at the end of the sequence. **B.** In source space, we found common differences between out- (top panel) and in-ward (bottom panel) conditions in left motor areas and right retrosplenial cortex. Specific to the inward condition, the posterior cingulate was more activated in the CTRL trials. * < .05, ** < .01

The contrast ILL versus CTRL emphasized the differences due to the postdictive mechanism (Fig. 3.14B). A significant increase in CTRL compared to ILL was found in the left motor areas after the sequence, for both direction conditions (outward: p = .044, 560-628ms; inward: p = .027, 420-420ms, p = .048, 588-664ms). Even if the motor response was the same, the reaction time was on average 123 ms shorter for CTRL than ILL and this delay could account for the difference in motor area. To test this hypothesis, we compared CTRL and VER trials (Fig. 3.15), whose difference in RT was also around 133 ms and thus should yield the same differences in neural activity in left motor areas. No differences were found in the left motor cortex, but activity in VER significantly increased in the right motor in outward condition (BA4-rh, p = .03, 524-612ms). Besides the change in hemisphere, this divergence showed an opposite pattern to the one between ILL and CTRL, suggesting that delays in RT are unlikely to fully explain the initial difference found between ILL and CTRL. The left motor areas thus could be implicated in the postdictive temporal reorganization of events.

In the inward condition, the left posterior cingulate was less activated by ILL trials than CTRL (p = .02, 400-468ms), and such a difference was also observed in outward condition although did not reach significance. The same pattern than for ILL and VER was observed in the right retrospenial cortex: in outward condition, the activity evoked by S2 was increasing in CTRL for outward (p = .024, 424-496ms), and was decreased in CTRL for inward condition (p = .047, 348-400ms). This suggests that the retrosplenial cortex could also be functionally relevant in postdiction and is additionally specific to the sequence direction.

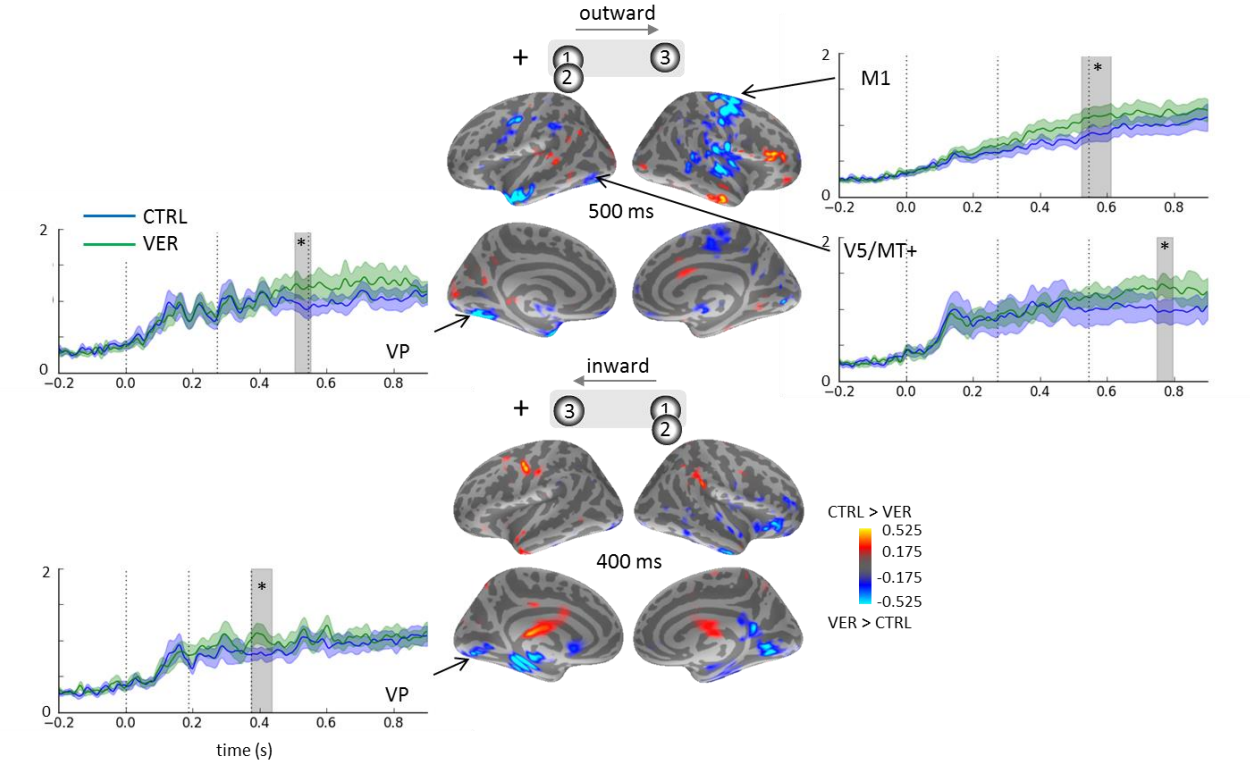


Fig 3.15 Contrasting difference stimulations and perceptions. The contrast CTRL versus VER was expected to yield to differences in the visual cortex contralateral to the stimulation, as the second stimulus was stimulated and perceived at a different spatial location. Indeed, for both in- and out-ward conditions, increased activity was found in the left ventral occipital areas in VER compared to CTRL around 150-200 ms after S2 onset (outward: p = .048, 504-552ms; inward: p =.036, 346-436ms). Specifically to the outward condition, the activity in the left MT+ (p = .04, 748-796ms) and in the right motor cortex (BA4-rh, p = .03, 524-612ms) significantly increased in VER relative to CTRL. *< .05, ** < .01. Abbreviations: BA: Brodmann area, VP: visual area VP, M1: primary motor area (BA 4), V5/MT: medial temporal, PCC: posterior cingulate cortex (BA 31), orbF: orbitofrontal (BA 10)

3.8.3. Spontaneous alpha power and phase

We hypothesized that increased uncertainty on sensory evidence would increase the weight of the "slow-speed" prior in the perceptual inference, resulting in a larger likelihood of reporting the rabbit illusion. As the power of ongoing alpha oscillations can be a marker of cortical excitability, we tested whether an increase of alpha power preceded an illusory perception. Since we targeted pre-stimulus activity, we combined inward and outward conditions. No significant effects between ILL and VER were found during pre-stimulus period (-600,-100ms) in sensor space or in whole-brain source estimate analyses. One possibility may be that the attentional strategy differs between in- and out-ward conditions, since the position of the first stimulus differs. Specifically, in the outward condition, the first stimulus is closer to the fovea whereas in the inward condition, correct performance would be obtained when the locus of spatial attention is directed more peripherally (Fig. 3.16B, diagrams). Hence, we considered the inward and outward conditions separately. In outward condition, an increase of pre-stimulus alpha power was found in three sensors of the right hemisphere between -590 and -490 ms (Fig. 3.16A). A source reconstruction of this significant temporal window showed that alpha power increased in the right visual cortices and left dorsal posterior cingulate when the illusion was perceived (Fig. 3.16B). These results suggest that the illusion is more likely perceived when spatial attention decreases as indexed by an increased alpha power. The dorsal posterior cingulate is also a generator of alpha rhythm within the DMN, implying that general attention can play a role in the perception of the illusion. However, in inward condition, no significant effects in pre-stimulus alpha power were found, suggesting a different attentional strategy potentially explained by the fact that the sequence started more peripherally.

We then tested the impact of the reliability of the first stimulus encoding on the processing of subsequent stimuli by analyzing the phase of spontaneous alpha oscillations. Both inward and outward conditions were combined. The alpha phase at the onset of the first stimulus was computed for each epoch at the sensor showing maximal pre-stimulus alpha power. For each individual, the phases were normalized relative to ILL. At the group level, each condition showed a preferential phase, as assessed by a Rayleigh test. However, a surrogate analysis showed in Fig. 3.16C that the two distributions were not significantly different (phase difference for mag: -0.26 rad, CI = [-0.79, -0.16]; grad: -0.37 rad, CI = [-0.98, -0.15]; eeg: +-0.12 rad, CI = [-0.46, 0.40]). We also performed the same analysis on the theta band centered on 7Hz, as in Chakravarthi & Van Rullen (2012). The two distributions for ILL and VER were not significantly different (phase difference for mag: -0.29 rad, CI = [-0.80, -0.12]; grad: -0.15 rad, CI = [-0.65, -0.03]; eeg: -0.24 rad, CI = [-0.40, 0.29]). The same analysis was conducted to investigate whether slow-frequency oscillations entrained by the visual sequence shape the sensory encoding. Results showed that the 4Hz phase did not distinguish illusory from veridical trials (phase difference for mag: -0.17 rad, CI = [-0.53, -0.02]; grad: -0.22 rad, CI = [-0.65, -0.06]; eeg: -0.01 rad, CI = [-0.06, 0.04]).

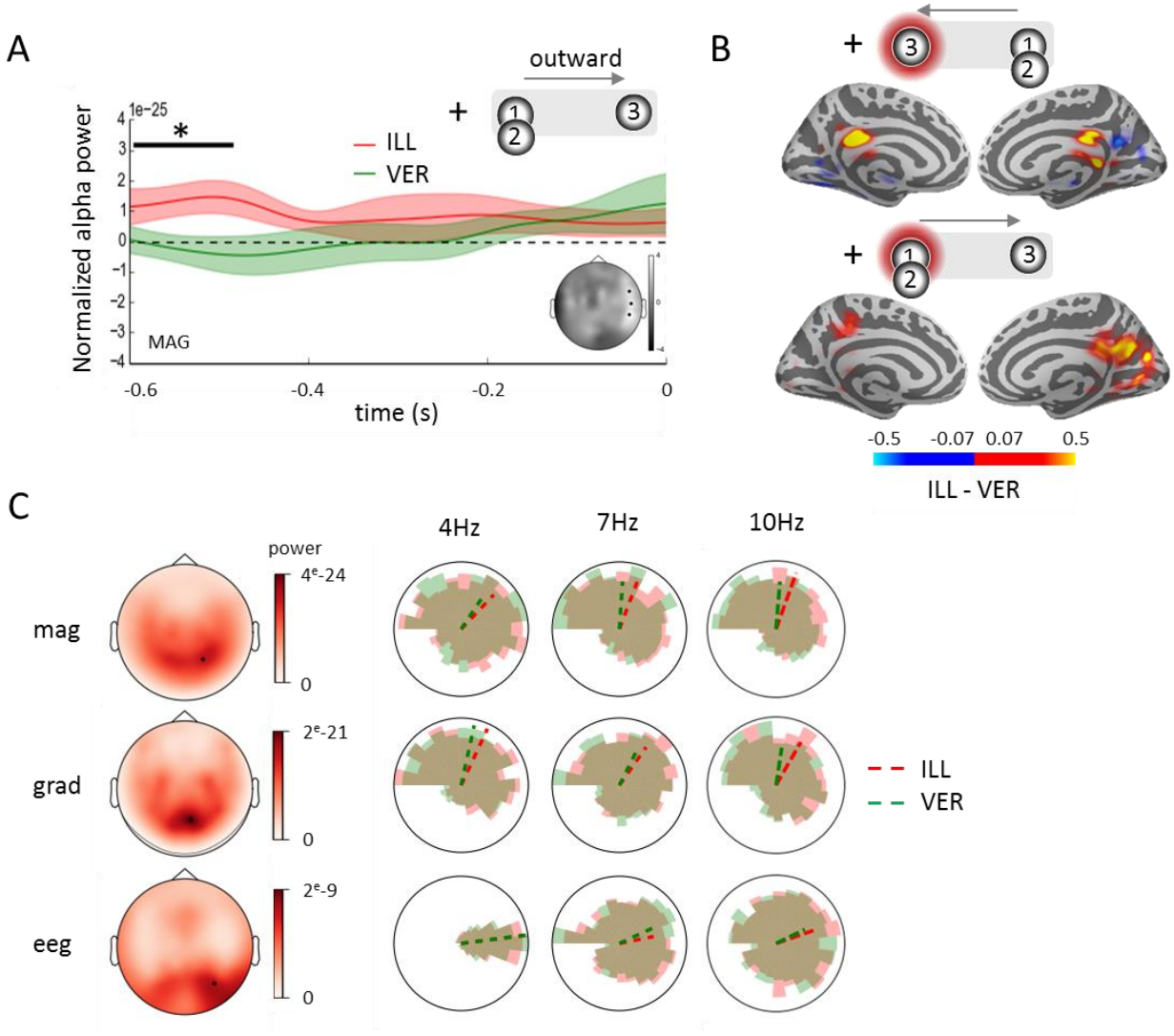


Fig 3.16. Prestimulus alpha power increases before the perception of the illusion only in outward condition. **A.** Spatio-temporal clustering between [-0.6,-0.1]s before stimulus onset showed significant clusters in sensors in outward condition. There are no significant clusters when outward and inward conditions are combined. **B.** This discrepancy could be explained by a change in attentional strategy: attention should be spatially focused closer to fovea in outward condition compared to the other condition (red circles). The sources showed an increased activation of the right visual cortex and both precuneus in ILL compared to VER conditions. **C**. The alpha phase at stimulus onset does not predict the perception of the illusion. One sensor showing maximal pre-stimulus slow-frequency power (2.4-13.3Hz, -0.6,-0.1s) when ILL and VER trials were averaged was selected for each sensor types (black dot on the topomaps on the left panel). For this sensor, the phase at stimulus onset was computed for each epoch and for each frequency band of interest (4, 7 and 10Hz). After normalization per individual, all phases were plotted on the right panel, for each sensor type. A Rayleigh test found significant preferential phase for ILL (red dotted line) and VER (green dotted line). However, the two conditions are not significantly different.

3.9. Discussion

3.9.1. Retrospective influence of top-down information on sequence regularity

Postdiction assumes that the perception of a visual sequence will derive from cognitive operations manipulating the information of the whole sequence (Blankenburg et al., 2006; Shimojo, 2014); hence, the spatial location of an earlier item may depend on information from the latest item in the sequence. In this experiment, perceiving the rabbit illusion was associated with increased activity in right parieto-frontal regions (orbitofrontal areas and right supramarginal gyrus) after the last item in the sequence. At the same time, the right retrosplenial/visual area showed an opposite pattern according to the direction: the amplitude increased when the illusion was perceived in inward trials, but decreased in outward trials (Fig. 3.13). Perceiving the illusion was also related with decreased activity in the left motor cortex, after the end of the sequence, when compared with CTRL trials (Fig. 3.14B).

The only other neuroimaging study investigating the rabbit illusion showed that the right premotor cortex was more activated during the illusion than during the correct perception of control sequences (Blankenburg et al., 2006). In the visual rabbit, we found increased activity in the left motor areas in CTRL as compared to ILL trials (Fig. 3.14B). The lateralization difference can stem from the stimulation side, on the right hemifield here, and on the left arm in the tactile rabbit, as the right hand was used in both cases as the motor effector. The increased activity in left motor areas was unlikely due to motor responses, since both reports were identical, neither due to the difference in RTs, since no such increase was observed in the contrast ILL versus VER with similar difference in RTs. Blankenburg and colleagues (2006) speculated that the premotor cortex was involved in top-down modulations of early somatosensory regions. The pre/supplementary motor area may indeed relate to motor but also non-motor explicit timing functions while the premotor cortex may be involved in temporal prediction (Coull & Nobre, 2008; see Coull et al., 2011 for a review). Here, the increased activity in CTRL found in the left motor cortex could translate a stronger temporal prediction due to the congruence of spatial and temporal cues (i.e. the sequence was regular in time and space in CTRL trials), compared to ILL, where the spatial and temporal cues were incongruent (i.e. the sequence was regular in time but not in space).

The right precentral inferior frontal gyrus (IFG) and middle prefrontal gyrus were also activated during the tactile rabbit illusion (Blankenburg et al., 2006). We found an increased activity in orbitofrontal areas (BA10) when the illusion was perceived but, given the coarse spatial resolution of MEG, the observed difference can be interpreted as in good agreement with the implication of IFG. The right IFG was largely reported as a generator of the mismatch negativity triggered by a deviant embedded in a regular sequence (Garrido et al., 2009). More interestingly, it was involved in the detection of regularity in auditory sequences (Barascud et al., 2016). The operation carried within the right IFG seems to retrospectively reassign to the sound sequence a regular temporal structure and to date a posteriori when the regularity began (Patel & Chait, 2011). In the rabbit illusion, the right IFG could provide top-

down prior information favoring spatial and temporal regularity in sequences. This is highly consistent with the activation in left IFG found when regularly spaced sequences, independently of physical stimuli, were contrasted with irregular spatial sequences (Blankenburg et al, 2006).

Alternatively, the orbitofrontal regions have been associated with conveying expectations in object recognition (Summerfield & Egner, 2009). Early visual inputs (~ 50 ms) could feed the orbitofrontal areas, returning back predictions about the nature of the object (Bar et al., 2006). Here, we observed an early increased activity, when the illusion is perceived, in outward condition at the second stimulus onset, meaning that the physical difference in stimuli between ILL and VER may have triggered varying predictions in the orbitofrontal areas. However, the difference arose later in the inward condition, which suggests an alternative explanation: the whole visual sequence could have been considered as a unique object. The retrosplenial cortex could be a possible candidate for receiving feedback predictions from the orbitofrontal area, as they are anatomically related by reciprocal connections (Vann et al., 2009). Here, the observed differences in the right retrosplenial cortex followed in both conditions the differences in orbitofrontal areas, which would be consistent with such an account. Besides, we found an opposite pattern of activation between direction conditions: in outward condition, perceiving the illusion was associated with a lower activity compared to a veridical perception (ILL or CTRL), while it was associated with a greater activity in inward condition. Noteworthy, the timing was slightly different for the two contrasts: the difference in ILL versus CTRL matched with the S2-evoked response while the difference in ILL versus VER was delayed to the end of the sequence. The delayed timing after the end of the sequence for this latter contrast speaks in favor of a postdictive reconstruction of the spatial location of the second stimulus. The right retrosplenial/lingual gyrus (BA30, 18 and 19) was found to be activated by biological motion and speed discrimination judgment (Orban et al., 1998), implying a role in the inference of biological form from motion information (Servos et al., 2002). This could explain the opposed pattern for centrifugal versus centripetal motion.

Last, we found increased activation in the supramarginal cortex (BA40) in illusory trials. The bilateral supramarginal gyrus (SMG) is involved in phonological loops tasks, when participants judge the sound of words (Hartwigsen et al., 2010). Right SMG is specifically implicated in temporal processing (Wiener et al., 2011), suggesting a potential role in the temporal structuring of information. The late timing of the observed difference (around 700-800ms in both conditions) suggests that the brain reaches a decision stage potentially based on the spatial feature of the target extracted from the sequence in the SMG.

3.9.2. Specificity of MT for outward motion

Left MT was more activated specifically in the outward condition in VER compared to CTRL and ILL. In other words, the MT activity increased when the second stimulus was perceived at the first position instead of the second, independently of the physical stimulation. MT neurons are known to be tuned to both direction and speed of moving stimuli (Perrone & Thiele, 2001; Born & Bradley, 2005; Liu & Newsome, 2005). A first tentative hypothesis could be that MT reacted to change in speed between S2

and S3 occurring in the contralateral visual field. The perceived speed between S2 and S3 should indeed be larger when S2 was perceived at the first position rather than at the intermediary one. An alternative hypothesis could be that MT is sensitive to speed changes within a sequence and would be more activated during temporally irregular sequence than regular ones. The specificity of the effect in the outward condition can be related to the fact that MT shows a bias toward centrifugal direction of motion (Albright, 1989), in the direction of the optic flow (Morrone et al., 2000). Moreover, half of the retinotopic MT area is devoted to the contralateral hemifield inside 15° (Born & Bradley, 2005). Difference in motion arising far outside this area, as in the inward condition, might generate a lower differential neural activity. Both the great eccentricity and the preferential bias for outward direction of motion might explain why, if such comparable effect was actually true for the inward condition, they might have been below the MEG sensitivity threshold.

3.9.3. Pre-stimulus spatial attention modulates the perception of illusion

For a Bayesian observer, increased uncertainty on the spatial location of a stimulus displaced the perceived position towards the prior distribution. Selective attention increases the gain of sensory signal in relevant areas while decreasing it in irrelevant areas (Jensen & Mazaheri, 2010). An increase in alpha activity is thus a possible marker of increased sensory uncertainty (Lange et al., 2013; Chaumon & Busch, 2014). We showed that increased alpha power in visual areas in the outward condition predicts the perception of the illusion (Fig. 3.16A). The fluctuations of spatial attention are likely to explain trial-by-trial performance. In outward condition, the to-be-judged second stimulus was presented closer to the fovea along with the first stimulus whereas in inward condition, the two first stimuli were presented more peripherally. Since detecting peripheral stimulus recruits more attentional resources (Tsal, 1983), participants might have preferred a more parsimonious strategy depending on the direction condition. Knowing that the instructions were to indicate whether the second stimulus was at an intermediary position or not, they could have compared the to-be-judged stimulus position with the position of the closest one to the fovea. In other words, in inward condition, they might have compared the position of the two last stimuli, while, in outward condition, they might have compared the two first stimuli. In that case, spontaneous fluctuations of attention occurring before the first stimulus would only influence perceptual outcomes in the outward condition. In inward condition, a bottom-up capture of attention could have been potentially triggered by the first stimulus (Buschman & Miller, 2007), which prepares the sensory system to efficiently process the following stimuli of interest, erasing any processing bias pertaining to pre-stimulus activity (Iemi et al., 2016).

Just as the influence of pre-stimulus alpha power appears limited to the processing of the two first stimulus of the sequence (S1 and S2), the impact of the alpha phase could have an even more limited temporal range (Palva & Palva, 2011). Alpha phase reflects the inhibition of local networks during brief interval and thus could determine the encoding reliability of S1 only. This would explain the absence of significant difference for the alpha/theta phase at the onset of the first stimulus between percepts. Alpha phase indeed reflects instantaneous fluctuations in cortical excitability of a local network (Ergenoglu et al., 2004; Romei et al., 2008), impacting probabilities for a visual stimulus to reach sensory

awareness (Busch et al., 2009; Mathewson et al., 2009). The phase at which S1 occurs, could be independent from the phase at which occurs S2, so would have no impact on the sensory encoding of S2. Our initial hypothesis (Fig. 3.10A) assumes that phases at S1 and S2 onset were related given that the rhythmicity of the sequence was entraining the neural oscillations. Neural entrainment provides a way to internalize external regularities (Schroder & Lakatos, 2009) in order to create sensory representation that can be manipulated (Lakatos et al, 2009). Since the phase of entrained oscillation was actively modulated to enhance perceptual performance (Stefanics et al., 2010; Henry & Obleser, 2012; Kösem et al., 2014), we hypothesized that the phase at S1 onset will influence the encoding and thus the perception of the whole visual sequence. Yet, no empirical evidence supports here this underlying assumption. An alternative can be that the first visual flash elicited a phase resetting of ongoing oscillations, making a fresh start for the oscillatory networks that are not biased anymore by previous ongoing activity (Makeig et al., 2002: Lakatos et al., 2007; Canavier, 2015). Phase-resetting can also explain the absence of results supporting the temporal window hypothesis (Fig. 3.10B), suggesting that an oscillation cycle was the neural implementation of a temporal window of integration shaping perception (Varela et al., 1981; Pöppel, 1997; Wutz & Melcher, 2014).

3.10. Further perspectives

It is important to note that the interpretations proposed in the *Discussion* section are temporary as analyses testing the oscillatory hypotheses are ongoing. For instance, more evidence is needed to support the statement that prior-based expectations from frontal regions feed into visual areas. Recent work suggest that top-down predictions updates are carried by alpha/beta oscillations and bottom-up surprise signals by gamma oscillations (van Kerkoerke et al., 2014; Michalaeras et al., 2016; Sedley et al., 2016). Alpha/beta increase in frontal areas and gamma modulation in visual areas could sign the dynamics of prior expectations in illusory trials. Alternatively, connectivity between frontal and visual areas could be enhanced within the beta range when the illusion is perceived. These analyses are currently under investigation.

Several theories suggest that the representation of sequence is underpinned by oscillatory coupling in working memory (Lisman & Jensen, 2013) or visual scene processing (Jensen et al., 2014). The temporal order of a sequence was supported by a theta/gamma neural code (Heusser et al., 2016), suggesting that the item reorganization occurring in the rabbit illusion may be unveiled by a difference in cross-frequency coupling. We are currently investigating these hypotheses.

Chapter 4

General discussion

4.1. Summary of empirical findings

The issues regarding the perception of the present moment reviewed in Chapter 1 confounds us into considering time both as a *dimension* that structures perception at large, and as a *conscious content* in the mind (van Wassenhove, 2017). Considering that neural oscillations have been proposed to provide a temporal architecture for cognition (Varela et al., 1981; Pöppel, 1997; VanRullen, 2016), my thesis work aimed at determining the intricate relationship between the temporal constraints on perception and temporal order perception itself; ultimately, I asked whether neural oscillations mechanistically encode perceptual moments in time.

In Chapter 2, we investigated how time may become an explicit conscious content by showing that temporal order is a psychological bias, which is stable over time and independent of attention. As foreseen by Pöppel (1997), the temporal window of successiveness, proposed as the building block of a temporal architecture in the mind, is a stable parameter that is specific to the anatomical and functional properties of each individual brain. This bias can be partially compensated for by attention due to the law of prior-entry, but it cannot be entirely obliterated. In a follow-up MEG study, we explored the impact of such intrinsic bias on the neural correlates of temporal order perception. A decrease in pre-stimulus alpha activity was found to predict the correctness of temporal order judgments, in line with the proposed gating-by-inhibition role of alpha oscillations (Jensen & Mazaheri, 2010). What was unexpected and novel in this study was the linear relationship of prestimulus alpha power fluctuations between an individual's perceived order and intrinsic bias. More precisely, participants having a large order bias (meaning that their perceived synchrony corresponded to a large physical asynchrony) also showed the largest fluctuations in ongoing alpha power when perceived orders were contrasted. Alpha oscillations have thus been interpreted as a means to compensate for an individual's internal bias, in line with the findings of the behavioral study. Additionally, we also tested the rhythmic temporal parsing hypothesis, but did found no evidence that pre-stimulus phase (in any frequency bands) predicted subsequent perceived orders. It could be that, in the absence of entrainment, the brain did not rely on a particular oscillatory phase to encode order.

In Chapter 3, we investigated how time implicitly shaped visual sequences. A striking consequence of discrete perception is that within the same temporal window of processing, late inputs may influence the representation of earlier information. We used the Rabbit illusion (Geldard & Sherrick, 1972) to study postdiction, in which the intermediate flash of a visual sequence is spatially mislocalized due to the temporal regularity of the sequence. We showed that when participants perceived the illusion, parieto-frontal regions were more activated following the presentation of the full visual sequence, suggesting that these high-order regions may carry top-down information about a prior favoring low-speed motion. The illusion was also partially modulated by attentional mechanisms and we found that pre-stimulus alpha power predicted the perception of the illusion only when the encoding of the first flash mattered in the decision. Both pre-stimulus, as attention, supported by alpha oscillations, and post-stimulus influence, as predictions based on the temporal structure of the stimulation,

combined to produce the Rabbit illusion. This dual interference between prediction and postdiction is probably a universal signature for perception.

In the present Chapter, I will first address the rhythmic temporal parsing hypothesis by contrasting our study showing negative results with recent investigations reporting positive results. Then, I will propose some hypotheses on the origin of the individual bias in temporal order perception and discuss the interplay between this bias and attention. This brings us to formulate a new working hypothesis: attention could be a mechanism to counteract intrinsic biases. The alpha power should therefore be re-examined as a marker of perceptual bias.

4.2. Does the phase of neural oscillations encode time?

We hypothesized that order could be coded by the phase of neural oscillations (Chapter 2, section 1.4) but we did not find any phase differences that could explain the perceived order of an audiovisual pair of stimuli (Chapter 2, section 3). In the meantime, other studies have also explored the idea of a phase code for perceived time and found positive results (Ikumi, 2016, Chapter 3, section 3.5; Milton & Pleydell-Pearce, 2016; Takahashi & Kitazawa, 2017; Fig. 4.1). These apparent contradictory results deserve some attention and the implication of variations in task, modality and techniques needs to be discussed (Table 1).

Study		Task		N	Results
Varela *et al.*, 1981	Visual SJ	- 3-AFC: simultaneous, motion, sequential - ITI unknown - Asynchrony of maximal perceptual uncertainty	EEG	5	- Occipito-parietal alpha phase (stimulus onset) predicts simultaneous *vs.* motion/sequential reports
Ikumi *et al.*, 2016	Audiovisual SJ	- ITI = 3.1s - Asynchrony of maximal perceptual uncertainty (method of constant stimuli)	EEG	10	- Opposite phase at 13 +/-2Hz (200ms pre-onset) predicts reports of simultaneity *vs.* non-simultaneity in AV trials, but not in VA trials - No effect of pre-stimulus alpha power
Milton *et al.*, 2016	Visual SJ	- ITI = 1.3s - Asynchrony of maximal perceptual uncertainty (staircase procedure)	EEG	20	- Opposite alpha phase? (7.8- 12.7Hz) (50 ms pre-onset) predicts reports of simultaneity *vs.* non-simultaneity perception - The phase effect was independent of alpha power changes triggered by attention.
Takahashi *et al.*, 2017	Tactile TOJ	- ITI = 2.3s - SOA = +/- 100ms (43-20% incorrect) - Crossed arms	MEG	16	- Parieto-occipital alpha phase (-70 to +110 ms post-onset) predicts correct *vs.* incorrect order - No effect for visual, auditory alpha or mu rhythm.

Table 1. Review of the literature investigating the hypothesis of a phase code for time perception.
3-AFC: three alternative forced choice, ITI: inter-trial interval, SOA: stimulus onset asynchrony, TOJ: temporal order judgment, SJ: simultaneity judgment.

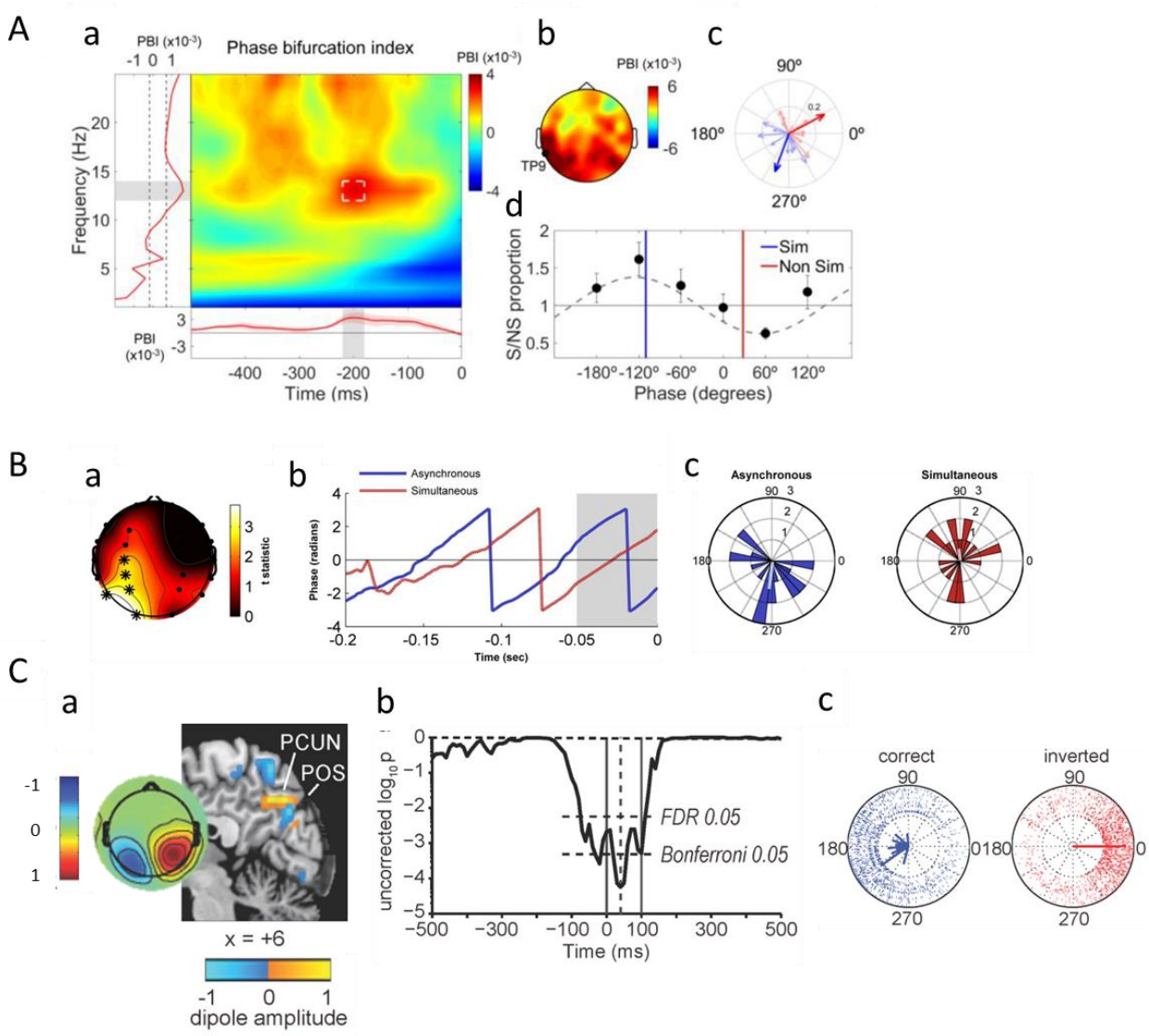


Fig. 4.1 The phase code hypothesis for time perception A. The bifurcation index (BI) averaged across channels and participants indicates that the phase between simultaneous and non-simultaneous audiovisual percept significantly differs (a, red) in AV trials. The maximal BI was found at 13Hz -200ms before stimulus onset in TP9 electrode (b). The mean phase angle for simultaneous (blue) are opposed to the phase angle for non-simultaneous percepts (c, red). Audiovisual perception fluctuates with the phase of pre-stimulus 13Hz oscillation (d). **B.** In a visual order judgment, the phase of alpha oscillation predicted perceived simultaneity (b). The analysis was conducted on the electrodes responsive to an attentional manipulation orthogonal to the order judgment (a). The mean phase angle for simultaneous trials (ligth red) is opposed to the phase angle for non-simultaneous percepts (c, light blue). **C.** Peri parieto-occipital alpha (a) predicted the perceived tactile order reversal triggered by arm crossing. A score computed from p-values indicates that the phase of the alpha component significantly differs between perceived orders (horizontal dashed line). Opposite alpha phase, after individual realignement relative to the preferential phase, predict correct (blue) or reversal order (red) perception (c). PCUN: precuneus, POS: parieto-occipital sulcus. (Adapted from: A: Ikumi, 2016; B: Milton et al., 2016; C: Takahashi & Kitazawa, 2017)

4.2.1. Theoretical and mechanistic ambiguity: integration *vs.* segregation

One first difference between studies reporting phase effects *vs*. those reporting no phase effect is that the former used simultaneity judgment (Ikumi, 2016; Milton & Pleydell-Pearce, 2016; but see Takahashi & Kitazawa, 2017, discussed later). Both theoretical and empirical differences distinguish SJ and TOJ tasks and it is in fact not surprising that the neural correlates of simultaneity and order perception differ.

The phase code hypothesis should be differently scrutinized depending on whether we consider integration or segregation. If the integration window is supported by neural oscillations, phase does not need to carry any temporal information, since the purpose of such window is to realign in time whatever falls within a cycle. Phase rather determines the boundary of the window, shaping the wax and wane of cortical excitability (Jensen & Mazaheri, 2010). Empirical evidences showing that near-to-opposite phase in the alpha band predicts simultaneity perception (Ikumi, 2016; Milton & Pleydell-Pearce, 2016), are consistent with this all-or-none account for phase (Fig. 4.1A and B). Consistent findings were also showed in studies using integration-based tasks such as tactile discrimination (Baumgarten et al., 2015) or the sound-induced double flash illusion (Cecere et al., 2015).

Specifically, integrating information implies that temporal order is lost, while segregating information aims at retrieving the temporal relationship between distinct events. Simultaneity consistency (Harris et al., 2010) is the precise purpose of numerous active processes compensating for neural delays, such as recalibration (Fujisaki et al., 2004; Vroomen et al., 2004), temporal ventriloquism (Spence & Squire, 2003) or distance-adjustment (Sugita & Suzuki, 2003). Recalibration was for instance observed both in simultaneity (Fujisaki et al., 2004) and order judgment (Vroomen et al., 2004) between simple sounds and flashes stimuli. However, when more ecological stimuli such as faces and speech were used, recalibration occurred in simultaneity but not anymore in order judgment (Vatakis et al., 2008). This discrepancy could be explained by the fact that recalibration is guided by contextual information (Roseboom & Derek, 2011). In lab situations, the system unnaturally implements compensatory processes such as recalibration in TOJ where they would not have been necessary with natural settings (Krakauer et al., 2017). Optimizing the perception of simultaneity ensures a correct detection of temporal coincidence, which is an important cue for second-order processes such as multisensory integration (Spence & Squire, 2003; King, 2005). While perceived simultaneity seems as a “first-order percept” used in additional computations, the status of order perception is apparently different. Perceiving order could be on the opposite considered as a “second-order percept” derived from perceived causality (Bechlivanidis & Lagnado, 2016), easier to instantiate in a naturalistic environment providing a rich knowledge on putative causal relationships (Buehner & May, 2003). This asymmetry between perceived simultaneity and order can explain why simultaneity and order judgment do not correlate (Love et al., 2013; van Eijk et al., 2008).

On the other hand, segregating events requires retaining temporal information. One hypothesis proposed that the phase of neural oscillations offers the temporal granularity necessary to encode

perceived order (see Chapter 2, section 1.4). Our failure to find such neural code in the perceptual domain suggests that order information might not be encoded in the phase even if it can be later recalled and handled via a phase code representation (Heusser et al., 2016). If order is indeed a "second-order percept" derived from causality, ongoing activity is more likely influencing first causality. Indeed, empirical evidence showed that the fronto-central alpha phase predicts causality perception (Cravo et al., 2015, Fig. 4.2). The phase opposition is nearly perfect (Fig. 4.2C), suggesting an all-or-none encoding similar to the one in simultaneity perception.

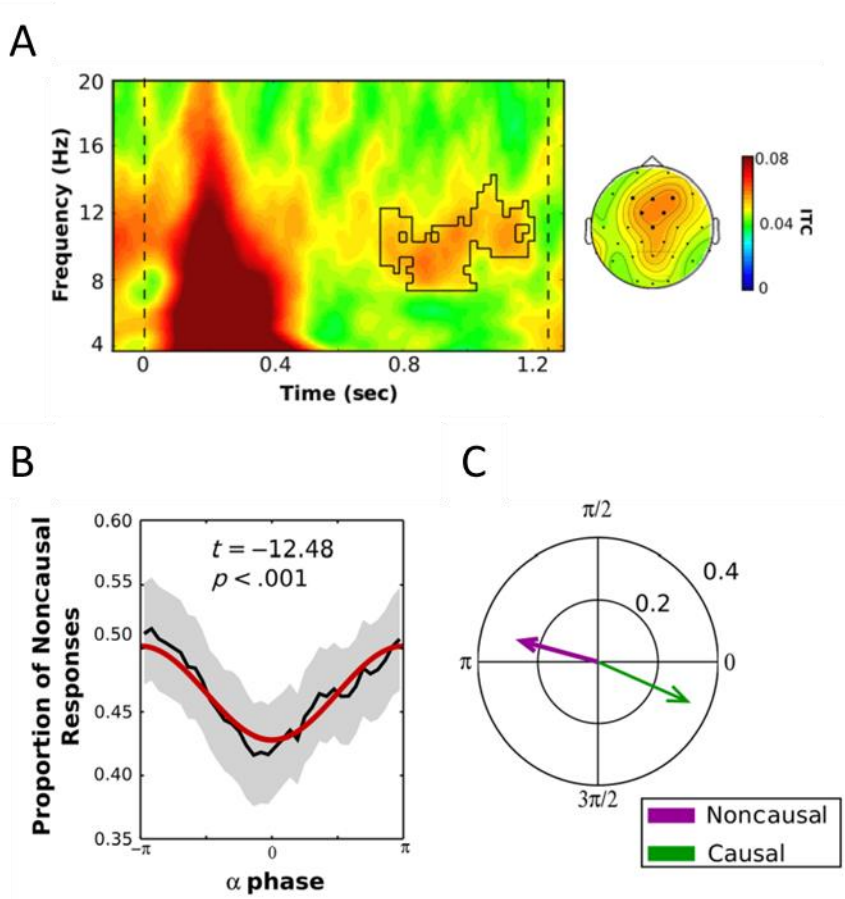


Fig. 4.2 Alpha phase predicts causality perception in the launching effect. A. The inter-trial coherence (ITC) is plotted for each time and frequency point. The dashed lines represent respectively the beginning of the motion of the ball and the time of impact with the second static ball. ITC in alpha band is significantly above chance level before the impact (black contours) in fronto-central electrodes (right inset, black). **B.** The proportion of non-causal responses is modulated by the alpha phase in the condition maximizing uncertainty. **C.** The mean alpha phase is opposed between causal (green) and non-causal (pink) judgment. (Adapted from Cravo et al., 2015).

However, one study seems to contradict this view as a phase difference between perceived tactile orders was reported (Takahashi & Kitazawa, 2017). Crucially, the incorrect perception of order was due to a mapping error caused by arm crossing. Without arm crossing and for the selected SOA, participants perceived the correct order in almost all the trials. This is significantly different from our

study and the simultaneity studies (Ikumi, 2016; Milton & Pleydell-Pearce, 2016) in which the temporal delay between stimuli was chosen as maximizing perceptual ambiguity. Second, the difference in phase was maximal 40ms post-onset, suggesting a second-stage involvement in the process. Third, the 10 Hz phase difference was mainly source localized in the bilateral precuneus, and not in the motor cortex as it would have been expected. Knowing that the precuneus is a central hub for allocentric spatial representation, the authors proposed that 10Hz oscillations in these regions could modulate information transfer. The early anatomic representation from the somatosensory cortices can be periodically suppressed or enhanced resulting in a correct or erroneous order judgment. To sum up, the 10Hz phase in this study does not code for perceived time so does not contradict the previous reasoning.

4.2.2. Sensory-dependency of time encoding: multisensory *vs.* unisensory

A second difference between these studies is the implicated modality: audiovisual in Ikumi and colleagues (2016), visual in Milton & Pleydell-Pearce (2016), and tactile in Takahashi & Kitazawa (2017). The ongoing phase was expected to modulate the excitability of the sensory area processing incoming stimulus, which is therefore easily predictable for unimodal judgment. In cross-modal judgment, both sensory areas involved can be potential candidate for a phase difference, as we hypothesized in our study. However, Ikumi and colleagues found that opposite phase of 13Hz oscillations predicted audiovisual simultaneity perception but only for pairs where the sound was presented first (Ikumi, 2016). This specificity could be explained by some modality-specific processes. Others studies related to timing also reported an asymmetry between vision and audition (Romei et al., 2007; Vidal, 2017; Cecere et al., 2017) or observed effects in the auditory cortex only (Kösem et al., 2014; Kaganovich & Schumaker, 2016), suggesting that the auditory cortices could be particularly suited for temporal processing (Zatorre & Belin, 2001).

4.2.3. Methodological differences: EEG *vs.* MEG

The last difference is methodological: the phase differences were assessed with EEG in Ikumi and colleagues, and Milton & Pleydell-Pearce, and with MEG in Takahashi & Kitazawa and in our study. A recent paper listed the studies that found phase-dependency in perception (VanRullen, 2016). This list, completed with recent studies, showed that 6 out of the 39 non-invasive studies used MEG, the others used EEG (Table 2). This imbalance could be due to the larger number of studies using EEG, a facilities much more widespread than MEG platforms. Another less trivial alternative would be that phase is more difficulty captured by MEG. Out of the 6 studies reporting significant phase differences in MEG, two used entrained oscillations (Kösem et al., 2014; Arnal et al., 2015), and three derived the phase from indirect measures: independent components (Takahashi & Kitazawa, 2017), gamma power envelope (Landau et al., 2015), phase difference index based on inter-trial coherence (Bonnefond & Jensen, 2012). Only Baumgarten and collegues found a significant phase difference with a direct measure and without benefiting from entrainment (Baumgarten, et al., 2015). While the picture is inevitably incomplete since the number of failures to report significant phase results remains unknown, a careful comparison would deserve to be conducted within one study combining MEG and EEG.

Category	Studies	Techniques
Threshold-level perception in visual modality	Nunn & Osselton, 1974	EEG
	Busch, Dubois, & VanRullen, 2009	EEG
	Mathewson, et al., 2009	EEG
	Busch & VanRullen, 2010	EEG
	Dugue, Marque, & VanRullen, 2011	EEG
	Fiebelkorn et al., 2013	EEG
	Hanslmayr, et al, 2013	EEG
	Neuling et al., 2012	EEG
Threshold-level perception in auditory modality	Rice & Hagstrom, 1989	EEG
	Ng, Schroeder, & Kayser, 2012	EEG
	Strauss, et al., 2015	EEG
Threshold-level perception in somatosensory modality	Ai & Ro, 2014	EEG
Supra-threshold perception	Callaway & Yeager,1960	EEG
	Dustman & Beck, 1965	?
	Drewes & VanRullen, 2011	EEG
Oculomotor functions	Hamm, Dyckman, McDowell, & Clementz, 2012	EEG
	McLelland, Lavergne, & VanRullen, 2014	EEG
Inattention and visual search	Buschman & Miller, 2009	ECoG in monkeys
	Busch &VanRullen, 2010	EEG
	Dugue, Marque, & VanRullen, 2015	EEG
	Landau, Schreyer, van Pelt, & Fries, 2015	MEG
	Voloh,Valiante, Everling, & Womelsdorf, 2015	ECoG in monkeys
Temporal parsing	Varela, Toro, John, & Schwartz, 1981	EEG
	Chakravarthi & VanRullen, 2012	EEG
	Cravo, Santos, Reyes, Caetano, & Claessens, 2015	EEG
	Baumgarten, Schnitzler, & Lange, 2015	MEG
	Kösem, et al., 2014	MEG
	Cecere et al., 2015	EEG
	Ikumi, 2016	EEG
	Milton & Pleydell-Pearce, 2016	EEG
	Ronconi & Marotti, 2017	EEG
	Takahashi & Kitazawa, 2017	MEG
Decision making	Wyart, de Gardelle, Scholl, & Summerfield, 2012	EEG
Top-down influence of predictions and expectations	Arnal, Doelling, & Poeppel, 2015	MEG
	Samaha, Bauer, Cimaroli, & Postle, 2015	EEG
	Ten Oever,van Atteveldt, & Sack, 2015	EEG
	Sherman, Kanai, Seth, & VanRullen, 2016	EEG
	Han & VanRullen, 2017	EEG
Cross-modal integration	van Erp,Philippi, de Winkel, & Werkhoven, 2014	EEG
Short-term memory	Siegel, Warden, & Miller, 2009	ECoG in monkeys
	Bonnefond & Jensen, 2012	MEG
	Myers, Stokes, Walther, & Nobre, 2014	EEG
	Leszczynski, Fell, & Axmacher, 2015	EEG

Table 2. Review of the studies which found phase-dependent effects, carried out by VanRullen, 2016. Recent and missing studies have been added in italics. Only 6 studies used MEG.

4.2.4. Summary

Contrary to predictions, we did not find any phase difference pertaining to time perception. However, this apparent discrepancy can be explained by variations in the tasks, electrophysiological techniques and stimuli modality used in the other studies (Ikumi, 2016, Milton & Pleydell-Pearce, 2016, Takahashi & Kitazawa, 2017). First, we argued that the perception of order could be a "second-order" percept derived from perceived causality, while the perception of simultaneity could be a "first-order" percept. Simultaneity and causality may be encoded by an all-or-none phase code. The mechanisms underlying order encoding remain unclear even if a fine-grained phase code might be involved in the maintenance and manipulation of memorized order. Second, audiovisual timing perception could also involve different mechanisms depending on the modality of the first physically presented stimulus. Third, it could be methodologically more challenging to detect phase differences with MEG than EEG, for some unknown reasons which would deserve some attention if this turns right.

4.3. New working hypothesis: attention as a mechanism to compensate for intrinsic biases

4.3.1. Origin of the individual bias in temporal order perception

We highlighted in Chapter 2.2 the existence of an individual bias in temporal order perception, which was reminiscent of an age-old debate on persistent errors found when astronomers estimated audiovisual delays to infer stars' speed (Sanford, 1888). It also potentially resonates with the Colavita effect, showing that observers preferentially rely on the visual modality when presented with audiovisual stimulation (Colavita, 1974). While this visual dominance appears strongly guided by attentional mechanisms, it could still originate from structural biases in the brain, since a complete reversal of the effect, i.e. an auditory dominance, was not obtained when manipulating attention (Sinnett et al., 2007). Yet, more investigations are needed to determine if the Colavita effect also shows inter-individual variability and if these two measures actually share the same origins. A recent study reported similar stable biases in temporal integration in task-specific manner (Ipser et al., 2017). Such persistent biases probably betray the temporal constraints grounded on the anatomic and functional connections of each individual brain, through which information is transferred.

One earlier study suggested that large inter-individual differences in temporal order perception arose from individual differences in neural delays (Stone et al., 2001). The time required to process sensory signals can vary depending on the attributes of the transmitted information: for instance, motion is processed faster than color (Moutoussis & Zeki, 1997). In neurosciences, neural latencies tend

to be implicitly assumed to equate subjective timing yet this assumption is highly debated (Moutoussis & Zeki, 1997; Arnold et al., 2001), was theoretically dismissed (Dennett & Kinsbourne, 1992; van Wassenhove, 2009) and received little empirical supports from the recalibration effect (Fujisaki et al., 2004; Vroomen et al., 2004) or neural correlates of prior-entry (McDonald et al., 2005). Instead, numerous active processes are implemented in the brain to compensate for these neural delays: active recalibration perceptually resynchronizes recurrent occurrences of asynchrony (Fujisaki et al., 2004; Vroomen et al., 2004); additional knowledge on the distance between stimuli (Sugita & Suzuki, 2003; Kopinska & Harris, 2004 ; Heron et al., 2007) or on causality (Bechlivanidis & Lagnado, 2016) guides perceived simultaneity and order; temporal ventriloquism (Spence & Squire, 2003) realign events to achieve simultaneity consistency (Harris et al., 2010).

Recently, neural delays were proposed to be compensated relative to the average timing across cognitive processes (Freeman et al., 2013). The "temporal renormalization" proposal was elaborated on the unexpected observation that individual PSSs in TOJ and the width of individual temporal windows of integration were anti-correlated (Freeman et al., 2013). The brain may thus estimate true external timing based on the distribution of neural latencies (Fig. 4.3A) and if this distribution is biased, perceived simultaneity will differ from veridical simultaneity (Fig. 4.3B). The bias in neural latency can arise from differences in the strength of structural connectivity between sensory and multimodal areas (Ipser, PhD thesis, p. 237). For instance, an interesting case study investigated the temporal segregation and integration abilities of a patient who consciously hears people speaks before seeing their lips move (Freeman et al., 2013). His PSS was indeed positive, indicating that the visual stimulus had to be presented first so that he could perceive audiovisual simultaneity. His impairment could originate from a general slowing of auditory processing due to his lesions in the inferior colliculus. One can imagine that such relative slowing can also arise in healthy participants due to slight differences in neural connectivity.

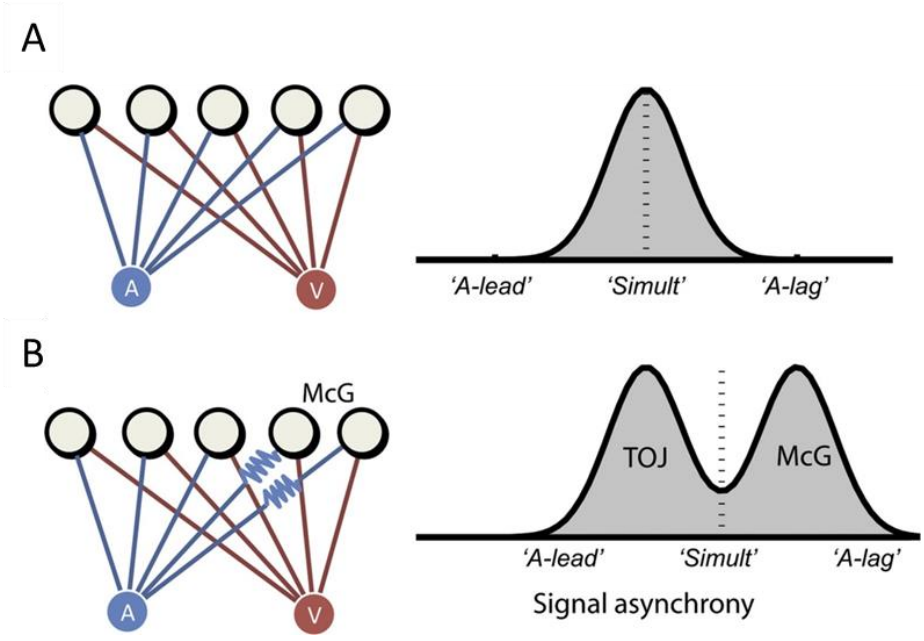


Fig. 4.3 Temporal renormalization theory. A. A synchronous pair of audiovisual stimulus (A and V) is processed following different neural pathways (circle). On the right, the distribution of neural asynchrony leads to different subjective report (x-axis). The mean of the distribution corresponds to the more probable timing for external timing, hence the report of subjective simultaneity (« simult »). **B.** If one pathway is particularly delayed (blue zigzag), for instance the auditory processing in a MacGurk (McG) task requiring audiovisual integration, the distribution of neural latencies across the different mechanisms becomes bimodal. The perceived asynchrony is renormalized relative to the mean of the new distribution, turning formerly « A-lag » into « simultaneous » percepts for unaffected tasks as a temporal order judgment (TOJ). (Adapted from Freeman et al., 2013).

Inter-individual differences in cognition have been related to differences in brain structure measured by Diffusion Tensor Imaging (Kanai & Rees, 2011). For instance, visual acuity and susceptibility to visual illusions were correlated with the size of retinotopic visual areas (Kanai & Rees, 2011) and grey matter volume in precentral regions predicted individual ability to learn intervals in the range of milliseconds (Bueti et al., 2012). The largest variability in functional connectivity between individuals was observed in multimodal association cortex (Fig. 4.4A), and was also related to sulcal depth but not cortical thickness (Mueller et al., 2013). These areas are also the ones supporting high-range connectivity (Fig. 4.4B), suggesting that large functional networks feed the cognitive plurality while local connectivity preserves modular core computations. The network of temporal order judgment, broadly the temporal parietal junction (inferior parietal cortices and supramarginal gyrus) and the middle frontal gyrus (see Chapter 2, Fig. 2.3), largely lies within these regions of high inter-individual variability in functional connectivity and cognitive abilities. Fiber tracts from the auditory cortex project to the superior temporal sulcus, intraparietal sulcus, supramarginal gyrus, and visual areas (Beer et al., 2011; Fig. 4.4C). The individual bias in temporal order perception can arise from individual imbalance in the structural connections between these areas.

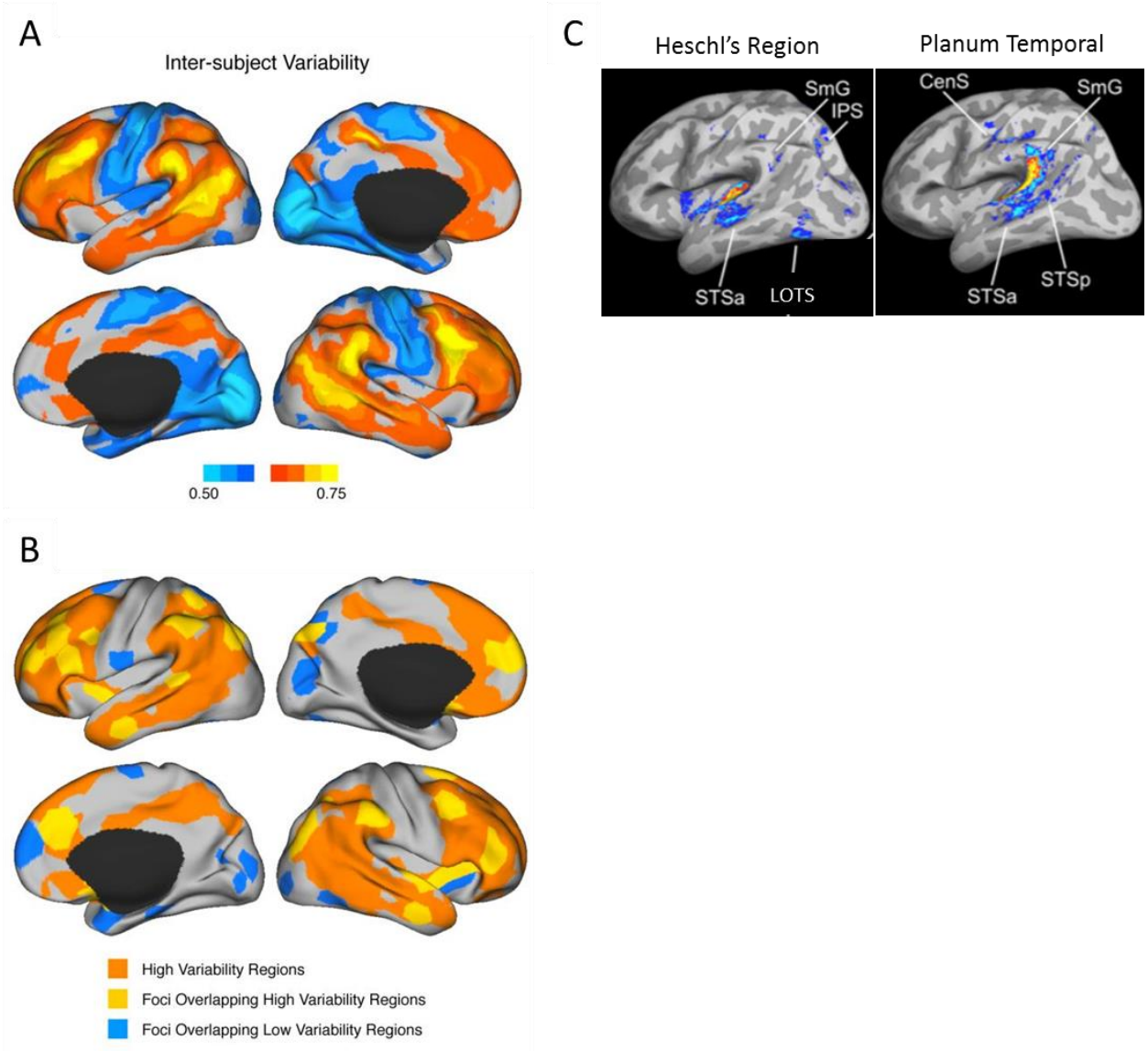


Fig. 4.4 Structural and functional connectivity of areas involved in the TOJ network. A. The association cortices showed the largest inter-subject variability in functional connectivity measured with fMRI during resting-state (yellow). **B.** Areas showing a correlation between inter-individual differences and functional connectivity are shown in orange. A meta-analysis was conducted of studies observing correlations between functional connectivity and individual differences in cognitive tasks. The resulting clusters overlap at 73% with regions of high variability in connectivity during resting-state (overlap shown in yellow). **C.** Cortical termination maps of tract fibers seeded in Heschl's region (left) and the planum temporale (right). Abbreviations: anterior and posterior superior temporal sulcus (STSa, STSp), central sulcus (CenS), supramarginal gyrus (SmG), intraparietal sulcus (IPS), lateral occipito-temporal sulcus (LOTS). (Adapted from: A, B: Mueller et al., 2013; C: Beer et al., 2011)

4.3.2. Interplay between structural bias and attention (as indexed by alpha oscillation)

In Chapter 2.3, we concluded that attention indexed by spontaneous alpha oscillations may partially counteract an individual temporal order bias. What is the interplay between such attentional compensation and the putative anatomical differences at the origin of the bias? Stronger connections between, for instance, visual areas and associative brain regions may generate a bias towards perceiving visual stimuli first in temporal order perception. Directing attention towards the auditory modality may inhibit visual processing and thereby increase the functional connectivity between the same associative regions and auditory areas, which would compensate for the initial imbalance (Fig. 4.5A). This hypothesis requires to test whether the strength of connectivity across implicated regions correlate with individual biases.

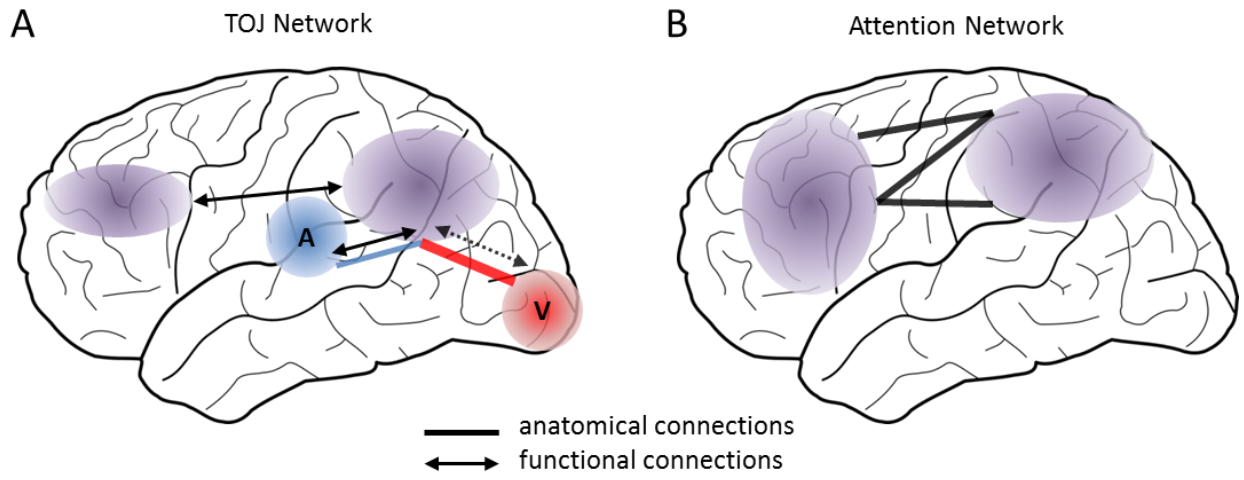


Fig. 4.5 Structural imbalance in the TOJ network or in the attentional networks. A. The individual bias in temporal order perception can arise from imbalanced anatomical connectivity between auditory (blue) and visual (red) areas. Anatomical connections between high-order areas (purple) can be stronger with visual areas (red thick line) than auditory areas (blue thin line). Attention can modulate the functional connectivity by inhibiting visual processing (dashed black arrow) thus enhancing auditory processing (solid black arrow). **B.** An alternative hypothesis is that the bias comes from anatomical differences in the attentional network, including, for instance, parietal areas related by the superior longitudinal fasciculus (black line) to frontal areas. Abbreviations: A: auditory areas, V: visual areas, TOJ: temporal order judgment. (B inspired from de Diego-Balaguer et al., 2016)

An alternative hypothesis would be that structural differences underlying an individual bias are not localized in the TOJ network but in an attentional network (Fig. 4.5B). For instance, the duration of stable percepts in bistable visual perception was negatively correlated with the cortical thickness of the bilateral superior parietal lobes (Kanai et al., 2010). The authors speculated that stronger thickness could entrain stronger feedback connections to early sensory areas, thereby triggering a faster

alternation rate between percepts. One could hypothesize that stronger top-down connections between one sensory area and associative regions in the TOJ network could facilitate the attentional capture of signals from this modality, hence favoring their processing. For instance, the visual attention network includes the ventral and dorsal superior longitudinal fasciculus and the inferior fronto-occipital fasciculus. A lateralization in the volume of the superior longitudinal fasciculus connecting the ventral and dorsal attentional pathway predicts larger spatial bias in an attentional task (Chechlacz et al., 2015). The microstructural properties of the parieto-frontal pathways also relate to inter-individual differences in attentional processing. Consistent with this, the cortical thickness in parieto-frontal areas was related to inter-individual differences in executive control and alerting components of attention (Westlye et al., 2011). The temporal parietal junction was also involved during spatial orienting but did not survive the correction for multiple comparisons. Knowing that this area also belongs to the TOJ network, one can speculate that the temporal parietal junction could be involved when attention is oriented toward one or another modality. Differences in this area's structure, being the cortical thickness or other features such as the grey/white matter density, could be at the origin of a bias in audiovisual temporal order.

The bias could thus arise from anatomical differences in the TOJ network, where stronger connections favor the processing of one modality compared to the other, or from differences in the attentional network, where stronger feedback connections modulate asymmetrically the processing of one modality (Fig. 4.5). In cognitive terms, the objective is to unravel whether the bias ensues from temporal order computations or from pure prior-entry. One way to disambiguate between these two hypotheses would be to test whether individual bias in audiovisual orienting correlates with the bias in temporal order perception, which would be in favor of the latter hypothesis. Another more direct experiment could compare with DTI which of the two networks correlates with individual biases in temporal order perception. This would bring some more general insights on how audiovisual processing and attention interface (Talsma et al., 2010).

4.3.3. Summary

The individual bias in temporal order perception, highlighted in this work, is probably one out of others biases deriving from anatomical and functional differences in brain structure. These structural differences could influence processing latencies of audiovisual stimulations but also all the numerous cortical mechanisms at stake to compensate for neural delays. The temporal renormalization theory (Freeman et al., 2013) proposes that the system compensates for neural latencies relative to an average estimate of neural delays across cognitive operations. Attention can partially counteract this bias, by modulating the functional connectivity between areas within the temporal order network, and thereby compensating for the imbalance of the anatomical connectivity. Alternatively, the bias could arise directly from differences in the strength of anatomical connections within attentional networks. More investigations are needed to disentangle these possibilities.

4.4. Alpha power as a state of biased perception

We propose that attention could compensate for intrinsic biases in temporal perception. This new working hypothesis has some direct implications on the functional role of alpha power.

4.4.1. Attentional fluctuations indexed by alpha power do not predict temporal order perception

A large body of literature investigated temporal order perception through the prism of the prior-entry, one of Titchener's laws (1908) stating that an attended stimulus will reach awareness sooner (Spence et al., 2001; MacDonald et al., 2005; Vibell et al., 2005; Weiss, Hilkenmeier, & Scharlau, 2013). Given that spontaneous alpha power potentially indexes attention (Jensen & Mazaheri, 2010; Klimesch, 2012; Sadaghiani & Kleinschmidt, 2016), a natural hypothesis derives from the law of prior-entry. When an audiovisual pair is presented, attending to audition should speed up the auditory processing and thus favor the perception of the sound first. Auditory alpha power should decrease to enhance auditory processing, while visual alpha power should increase to inhibit the distractor flow of visual inputs (Klimesch, Sauseng et al., 2005). On the contrary, attending to vision should favor the perception of the flash first and trigger a decrease of visual alpha power. However, we found no evidence supporting this hypothesis, and other studies also reported null effects of alpha power on audiovisual (Yuan et al., 2016), visual (Milton et al., 2016) and auditory (Bernasconi et al., 2011) simultaneity perception. That means that at least one of the two premises should be wrong: either the prior-entry is not at the root of temporal order perception, or spontaneous alpha power does not (only) index attention. Concerning the former, our first study indeed suggests that the point of subjective simultaneity (PSS) is independent of attention, and that the prior-entry might only explain perceptual deviations from the PSS (Grabot & van Wassenhove, 2017). Concerning the latter alternative, recent studies have challenged the general idea that attention-mediated alpha is enhancing perceptual sensitivity.

4.4.2. Alpha power as a state of biased perception

Spontaneous alpha rhythm has been usually negatively correlated with performance (Ergenoglu et al., 2004; van Dijk et al., 2008; Busch *et al.* 2009; Mathewson *et al.* 2009, Romei et al., 2010). However, some studies have empirically showed that low alpha power modulates the criterion but not the sensitivity of the perceptual decision (Lange et al., 2013; Limbach & Corballis, 2016; Iemi et al., 2016). Iemi and colleagues also conducted a literature review and concluded that the majority of studies showing a relationship between alpha power and performance had assessed performance with a criterion-dependent measure, as the hit rate, and conversely, most of the studies showing no effect were using sensitivity-dependent measures (also observed in Sherman et al., 2016). They argued that ongoing alpha oscillation reflects a state of biased perception where the probability of perceiving a stimulus is more or less likely, regardless of the actual stimulus presence.

Two scenarios could be envisioned where alpha would not entail the same functional implications. In the first case, as evidenced in the previously reviewed studies, alpha indexes a global cortical excitability modulation and does not improve sensitivity. It rather would act as a perceptual bias modulating both signal and noise. In the second case, alpha power can be decreased in specific cortical regions, thus lowering the criterion for relevant information and decreased in goal-irrelevant areas, thereby inhibiting distractor information processed in others areas (Jensen & Mazaheri, 2010). The differential modulation of power between cortical regions confers on alpha oscillations a functional selectivity. Studies on spatial attention showing decreased alpha in the contralateral relative to the ipsilateral hemisphere are consistent with this scenario (e.g. Thut et al., 2006), where sensitivity is effectively enhanced. In our study on temporal order judgment, we formulated a prediction in regard with these second-case scenario - namely that both high auditory alpha and low visual alpha should predict the perception of the flash first (and vice-versa). However, the results invalidated this prediction, which could suggest that our study actually fell under the first picture. I proposed that the alpha power biases perception in favor of an intrinsic hidden variable.

4.4.3. Summary

To summarize, our empirical findings does not support the hypothesis that temporal order perception relies on attentional modulations based on the law of prior-entry. They are more consistent with the hypothesis that spontaneous alpha rhythm reflects a perceptual bias favoring a decision irrespectively of the sensory evidence. In temporal order perception, the modulation of the alpha power reflects the likelihood of a participant to decide for its internal bias (the incorrect order) or for the correct order.

4.5. Conclusions and future perspectives for an individual approach to consciousness

In this work, we investigated temporal order as conscious content by tackling inter-individual differences. This approach has proved useful and particularly powerful to provide novel perspectives on temporal order perception but also on the functional role of alpha oscillations. Investigating inter-individual differences would probably provide similarly fruitful insights if we extend it to others domains investigating consciousness (Kanai & Rees, 2011; Dubois, 2016).

In cognitive neuroscience, objective reports are used as a basis to categorize mental states and to distinguish underlying neural mechanisms. One can argue that objective reports are not enough to describe the subtleties of mental states and therefore their underlying mechanisms (Lutz & Thompson, 2003). Indeed, perceiving a flash first does not correspond to the same neural operations in a

participant's brain showing an auditory bias compared to another participant's brain with a visual bias. Some even called for rigorous and extensive first-person reports in order to correlate them carefully with neural mechanisms (Lutz & Thompson, 2003). Even if group averaging procedure appears as a well-established method to sketch out the general principles of a phenomenon, some more elaborated processes can be missed in doing so (Matthews & Meck, 2014; Benwell et al., 2015). Inter-individual differences also inform on structural differences (Marshall et al., 2015; Chechlacz et al., 2015), which makes it an ideal approach to link brain, mind and neural mechanisms.

Appendix

During the last year of my PhD, I was part of a collaboration with Alexandre Gramfort (Télécom ParisTech, University Paris-Saclay, LTCI), Valérie Doyère NeuroPSI (University Paris-Saclay, Cognition & Behavior), Virginie van Wassenhove (NeuroSpin, Unicog), Tom Dupré La Tour, Tadeusz Kononowicz and Lucille Tallot. The project investigated cross-frequency coupling in time production, in both rats and humans, and aimed at developing an open-source cross-frequency pipeline. I was in charge of analyzing a set of MEG data where participants had to produce a fixed duration and to judge the accuracy of their own performance. We are currently finalizing the analyses and writing the paper (abstract in Appendix 1). A recently submitted methodological paper proposed a novel method to compute cross-frequency coupling avoiding some pitfalls of current methods (abstract in Appendix 2).

Appendix 1:
Cross-frequency coupling in time production

Grabot, L., Kononowicz, T., Dupré la Tour, T., van Wassenhove, V. (in prep) Cross-frequency coupling in time production.

Abstract

Humans can both estimate a duration and evaluate their self-performance. The exact neurocognitive mechanisms underlying these abilities remain elusive. One the one hand, recent evidence has involved beta power during time estimation (Kononowicz et al., 2015; Kulashekhar et al, 2016) and on the other, temporal attention is indexed by the phase of alpha oscillations (Rohenkohl et al., 2011) and temporal expectation by the phase-amplitude-coupling (PAC) of theta/beta (Cravo et al., 2011). Additionally, recent studies have shown that alpha/gamma PAC plays a role in the allocation of attention in visual tasks (Voyteck et al., 2010; Roux & Uhlhaas, 2013), and theta/gamma PAC has been proposed as a mechanism for temporal expectation (Henry & Herrmann, 2014). Here, we explored the hypothesis that alpha/beta PAC may be implicated in a task requiring time production and temporal self-estimation (i.e. attention to one's time performance). We exhaustively tested all possible functional oscillatory nesting in a time production task in which participants produced 1.45 s time intervals by pressing a button twice followed by self-evaluation on their temporal performances. PAC (Tort et al., 2009) was computed during time production, i.e. between the two button presses. PAC was also tested as control before the first button press, to test for motor preparation. A robust and significant alpha/beta PAC was readily observed for each individual (n=12). The alpha/beta coupling was specific to the timing production task but the coupling strength did not predict performance or the self-estimated duration. We will discuss the role of PAC in time estimation in the context of current proposals for temporal multiplexing (Gu et al., 2016) and the more general idea of the maintenance of current sensorimotor state (Engel & Fries, 2010).

References

Kononowicz, Tadeusz W., van Rijn, Hedderik. 2015. Neuropsychologia, 75, 381-389.

Kulashekhar, Shrikanth, Pekkola, Johanna, Palva, Jaakko Matias, Palva, Satu. 2016. 37(9) 3262–3281.

Rohenkohl, Gustavo & Nobre, Anna C. 2011. J Neurosci. 31, 14076-84.

Cravo, André M, Rohenkohl, Gustavo, Wyart, Valentin, Nobre, Anna C. 2011. J Neurophysiol. 106(6), 2964-72

Voytek, Bradley, Canolty, Ryan T., Shestyuk, Avgusta, Crone, Nathan E. , Parvizi, Josef, Knight, Robert T., 2010. Front Hum Neurosci. 4(191)

Roux, Frederic & Uhlhaas, Peter J, 2013, Trends Cogn Sci., 1-10

Henry, Molly J. & Herrmann, Björn, 2014, Timing & Time Perception, 2(1), 62-86

Tort, Adriano B. L. , Komorowski, Robert W., Manns, Joseph R., Kopell, Nancy J., Eichenbaum, Howard. 2009 PNAS 106(49), 20942-947

Engel, Andreas K., & Fries, Pascal. 2010, Curr Opin Neurobiol. 20(2):156-65.

Appendix 2:
Non-linear auto-regressive models for cross-frequency coupling in neural time series

Dupré la Tour T., Tallot L., **Grabot L.**, Doyère V., van Wassenhove V., Grenier Y., Gramfort A., *(submitted) Non-linear auto-regressive models for cross-frequency coupling in neural time series.*

Abstract

We address the issue of reliably detecting and quantifying cross-frequency coupling (CFC) in neural time series. Based on non-linear auto-regressive models, the proposed method provides a generative and parametric model of the time-varying spectral content of the signals. As this method models the entire spectrum simultaneously, it avoids the pitfalls related to incorrect filtering or the use of the Hilbert transform on wide-band signals. As the model is probabilistic, it also provides a score of the model ``goodness of fit'' via the likelihood, enabling easy and legitimate model selection and parameter comparison; this data-driven feature is unique to our model-based approach. Using three datasets obtained with invasive electrophysiological recordings in humans and rodents, we demonstrate that these models are able to replicate previous results obtained with other metrics, but also reveal new insights such as the influence of the amplitude of the slow oscillation. Using simulations, we demonstrate that our parametric method can reveal neural couplings with shorter signals than non-parametric methods. We also show how the likelihood can be used to find optimal filtering parameters, suggesting new properties on the spectrum of the driving signal, but also to estimate the optimal delay between the coupled signals, enabling a directionality estimation in the coupling.

Non-linear Auto-Regressive Models for Cross-Frequency Coupling in Neural Time Series

Tom Dupré la Tour[1*], Lucille Tallot[2,3], Laetitia Grabot[4,5], Valérie Doyère[3], Virginie van Wassenhove[4,5], Yves Grenier[1], Alexandre Gramfort[1,5,6]

1 LTCI, Telecom ParisTech, Université Paris-Saclay, Paris, France
2 Neuroscience Paris Seine, CNRS, INSERM, Sorbonne Universités, UPMC, Paris, France
3 Neuro-PSI, Université Paris-Sud, Université Paris Saclay, CNRS, Orsay, France
4 Cognitive Neuroimaging Unit, CEA/DRF/Joliot, INSERM, Université Paris-Sud, Université Paris-Saclay, NeuroSpin center, Gif-sur-Yvette, France
5 CEA, Université Paris-Saclay, Gif-sur-Yvette, France
6 Inria, Université Paris-Saclay, Saclay, France

* tom.duprelatour@telecom-paristech.fr

Abstract

We address the issue of reliably detecting and quantifying cross-frequency coupling (CFC) in neural time series. Based on non-linear auto-regressive models, the proposed method provides a generative and parametric model of the time-varying spectral content of the signals. As this method models the entire spectrum simultaneously, it avoids the pitfalls related to incorrect filtering or the use of the Hilbert transform on wide-band signals. As the model is probabilistic, it also provides a score of the model "goodness of fit" via the likelihood, enabling easy and legitimate model selection and parameter comparison; this data-driven feature is unique to our model-based approach. Using three datasets obtained with invasive electrophysiological recordings in humans and rodents, we demonstrate that these models are able to replicate previous results obtained with other metrics, but also reveal new insights such as the influence of the amplitude of the slow oscillation. Using simulations we demonstrate that our parametric method can reveal neural couplings with shorter signals than non-parametric methods. We also show how the likelihood can be used to find optimal filtering parameters, suggesting new properties on the spectrum of the driving signal, but also to estimate the optimal delay between the coupled signals, enabling a directionality estimation in the coupling.

Author Summary

Neural oscillations synchronize information across brain areas at various anatomical and temporal scales. Of particular relevance here, slow fluctuations of brain activity have been shown to affect high frequency neural activity by regulating the excitability level of neural populations. Such cross-frequency-coupling can take several forms. The most frequently observed type is when the power of high frequency activity is time-locked to a specific phase of slow frequency oscillations, yielding phase-amplitude-coupling (PAC). Even when readily observed in the neural recordings, such non-linear coupling is particularly challenging to formally characterize. Typically, neuroscientists use custom band-pass filtering or Hilbert transforms with cross-frequency correlations. Here, we explicitly address current limitations and propose an alternative probabilistic signal modeling approach for which statistical inference is fast and well-posed. To statistically model PAC, we propose the use of non-linear auto-regressive models which estimate the spectral modulation of a signal conditionally to a driving signal. This conditional spectral analysis enables easy model selection and clear hypothesis-testing by using the likelihood of a given model. We demonstrate the advantage of the model-based approach on three datasets acquired in rats and in humans. We further provide novel neuroscientific insights on previously reported PAC phenomena capturing two possible mechanisms for PAC: influence of amplitude and directionality estimation.

Introduction

The characterization of neural oscillations, which are observed in the mammalian brain at different temporal and spatial scales, have given rise to important mechanistic hypotheses regarding their functional role in neurosciences (e.g. [Buzsáki, 2006, Fries, 2015]). One working hypothesis suggests that the coupling across neural oscillations may regulate and synchronize multi-scale communication of neural information within and across neural ensembles [Buzsáki, 2010, Fries, 2015]. The coupling across different oscillatory activity is generically called cross-frequency-coupling (CFC) and has started to receive much attention [Jensen and Colgin, 2007, Lisman and Jensen, 2013, Canolty et al., 2006, Canolty and Knight, 2010, Hyafil et al., 2015]. The most frequent instance of CFC consists in the observation that the power of high frequency activity is modulated by fluctuations of low-frequency oscillations, resulting in phase-amplitude-coupling (PAC). Other instances of CFC include phase-phase coupling [Tort et al., 2007, Malerba and Kopell, 2013], amplitude-amplitude coupling [Bruns and Eckhorn, 2004, Shirvalkar et al., 2010], and phase-frequency coupling [Jensen and Colgin, 2007, Hyafil et al., 2015]. By far, PAC is the most reported CFC in the literature.

Seminally, PAC was described in local field potential (LFP) of rodents displaying a modulation of gamma band power (40−100 Hz) as a function of the phase of their hippocampal theta band (5−10 Hz) [Bragin et al., 1995, Tort et al., 2008]. Parallel recordings in different brain areas in behaving rodents have also highlighted differences in PAC between brain areas (e.g., hippocampus and striatum) at specific moments during a goal-oriented behavior, both in terms of which high-frequency range and how narrow-band the low frequency is [Tort et al., 2008]. PAC may promote cellular plasticity underlying memory formation [Axmacher et al., 2006]. In humans, theta (4−8 Hz)/gamma (80−150 Hz) PAC was described in human auditory cortex during speech perception [Canolty et al., 2006]. In more recent work, theta/gamma PAC was reported during the processing of auditory sequences in both humans and monkeys [Kikuchi et al., 2017], during working memory maintenance in human hippocampus [Axmacher et al., 2010], and during serial memory recall using non-invasive human magnetoencephalography (MEG) [Heusser et al., 2016]. PAC has been proposed to support the maintenance of information and to play an important role in long distance communication between different neural populations, considering that slow oscillations can propagate at larger scales than fast ones [Jensen and Colgin, 2007, Khan et al., 2013, Lisman and Jensen, 2013, Hyafil et al., 2015, Bonnefond et al., 2017]. Consistent with this notion, PAC has also been reported across distinct brain regions [Sweeney-Reed et al., 2014]. In sum, PAC has been proposed as a canonical mechanism for neural syntax [Buzsáki, 2010].

Given the growing interest in CFC, and in PAC more specifically, developing adequate and unbiased tools to quantify the posited signatures of neural computations has motivated a number of contributions [Canolty et al., 2006, Penny et al., 2008, Tort et al., 2010]. While these works have met some success, they have also pointed out numerous methodological challenges [Aru et al., 2015, van Wijk et al., 2015] illustrated and discussed in several simulation studies [Tort et al., 2010, Özkurt and Schnitzler, 2011]. Some important limiting factors and potential pitfalls of current used techniques are related to the impact of incorrectly choosing the bandwidth of bandpass filters [Dvorak and Fenton, 2014, Aru et al., 2015], the consequences of applying Hilbert transform on wide-band signals [Chavez et al., 2006, Dvorak and Fenton, 2014], and the potential misidentification of CFC [Hyafil, 2015, Kramer et al., 2008]. The direction we advocate to address these issues is to use a signal model that would fit the data in the sense of the mean squared error, and from which a measure of PAC or CFC could be derived. What all existing metrics have in common is that they do not make use of a signal model, and therefore setting the parameters of the method, such as filtering parameters, can only be driven by how much they lead to a strong PAC/CFC metric. As a consequence, even though current metrics give reasonable estimates of PAC, a legitimate and controlled comparison of methods and parameters, and therefore of the results, is impossible. Additionally, while simulations provide better control, they do not fully solve this issue faced by experimentalists since a simulation may approximate at best, or miss at worst, the real structure of neurophysiological signals. Hence, with the present contribution, we hope to propose a major improvement in CFC/PAC estimation by adopting a principled modeling approach.

To initiate this statistical model-based approach, we propose to consider a signal model which is rich enough to capture time-varying phenomena such as PAC. A model-based approach allows computing the likelihood of a recorded neural signal, which can be interpreted as a measure of the goodness of fit of the

model. In the present case, the goodness of fit corresponds to the classical measure of explained variance. Such an evaluation metric is a natural criterion to compare models, and a first step towards an automatic model parameters selection on brain data. Additionally, with a model-based approach, it becomes possible to answer a large variety of practical questions: From a statistical standpoint, one can define which parametrization or which hyper-parameter should be selected, while from a neuroscientific standpoint, one could ask, for instance, whether the instantaneous amplitude of slow oscillations contributes to PAC, or whether the amplitude modulation affects the entire neural spectrum identically. Importantly, our statistical signal modeling approach is not a biophysical one and is thus distinct from previous work [Hyafil et al., 2015, Chehelcheraghi et al., 2017]. Our goal here is to better explain and describe the empirical data themselves in the absence of any assumption regarding the neural mechanisms that have generated them.

To capture PAC in univariate time-series, we propose to use auto-regressive (AR) models, which are stochastic signal models as opposed to deterministic. A canonical deterministic signal model consists in modeling a time-series as a pure sinusoid corrupted by some additive noise. As recently discussed, considering neural oscillations as pure sinusoidal brain responses is an oversimplification of the complex neurophysiological reality [Cole and Voytek, 2017]. AR models do not make the strong assumption of sinusoidality and are rather based on the ability of the model to forecast the signal values. AR models have been successfully used to address multiple problems in neurophysiology such as spectral estimation [Spyers-Ashby et al., 1998], temporal whitening [Mahan et al., 2015], and connectivity measures like Granger causality [Granger, 1988, Valdés-Sosa et al., 2005, Haufe et al., 2010] or coherence [Baccalá and Sameshima, 2001].

In the present work, we show how driven auto-regressive (DAR) models [Grenier, 2013, Dupré la Tour et al., 2017] can be extended in order to study PAC. Standard AR models are statistically efficient given their low number of parameters, but they are linear, and therefore they cannot directly model non-linear phenomena like PAC. In order to extend AR models to cope with such situations of non-linearity and non-stationarity in signals, various advanced AR models have been proposed in other research fields such as audio signal processing and econometrics [Tong and Lim, 1980, Chan and Tong, 1986, Grenier and Omnes-Chevalier, 1988] (see related non-linear AR models section below). Here, we consider the slow oscillation as the exogenous driver so as to allow the coefficients of the AR model to be time varying, thereby making the instantaneous power spectral density (PSD) of the signal be a function of a slow exogenous time series. By doing so, DAR models are statistical signal models (and not biophysical ones) that do not model PAC explicitly but which are able to capture it.

In what follows, we first present how PAC is commonly analyzed in neural time series while pointing out limitations. We then review the current literature on non-linear AR models beyond neuroscience. We then detail the methodological aspects of DAR models first generically and then specifically for brain data. We describe how we can derive from a DAR model a power spectral density parametrized by an external signal and how it allows to compute so-called comodulograms and estimate directionality of coupling; this enables a fine-grained analysis of spectral modulations observed in PAC. Finally, results show how the comparison of goodness of fit between different models allow to conclude on optimal parameters (e.g. filtering parameters), leading to new neuroscientific insights, such as the wideband or the asymetric spectral properties of the driver. We also report that taking into account the instantaneous amplitude of the driver improves the goodness of fit, which suggests that this information is relevant to the coupling phenomenon and should not be discarded as in most PAC metrics. Last but not least, directionality estimation results reveal some delays between high frequency bursts and the driving low frequency oscillation.

Notations

We note y the signal containing the high-frequency activity, and x the signal with slow frequency oscillations, also called the exogenous driver. When a signal x results from a band-pass filtering step, we note the central frequency of the filter f_x and the bandwidth Δf_x. The value of the signal x at time t is denoted $x(t)$.

PAC quantification methods and current limitations

To estimate PAC, the typical pipeline reported in the literature consists in four main processing steps:

1. Bandpass filtering is performed in order to extract the narrow-band neural oscillations at both low frequency f_x and high frequency f_y;
2. A Hilbert transform is applied to get the complex-valued analytic signals of x and y;
3. The phase ϕ_x of the low frequency oscillation and the amplitude a_y of the high frequency signals are extracted from the complex-valued signals;
4. A dedicated approach is used to quantify the correlation between the phase ϕ_x and the amplitude a_y signals.

The Modulation Index (MI) described in the pioneering work of [Canolty et al., 2006] is the mean over time of the composite signal $z = a_y e^{\phi_x}$. The stronger the coupling between ϕ_x and a_y, the more the MI deviates from zero. This index has been further improved by Ozkurt *et al.* with a simple normalization [Özkurt and Schnitzler, 2011]. Another approach [Lakatos et al., 2005, Tort et al., 2010] has been to partition $[0, 2\pi]$ into smaller intervals to get the time points t when $\phi_x(t)$ is within each interval, and to compute the mean of $a_y(t)$ on these time points. PAC was then quantified by looking at how much the distribution of a_y differs from uniformity with respect to ϕ_x. For instance, a simple height ratio [Lakatos et al., 2005], or a Kullback-Leibler divergence as proposed by Tort *et al.* [Tort et al., 2010], can be computed between the estimated distribution and the uniform distribution. Alternatively, it was proposed in [Bruns and Eckhorn, 2004] to use direct correlation between x and a_y. As this method yielded artificially weaker coupling values when the maximum amplitude a_y was not exactly on the peaks or troughs of x, this method was later extended to generalized linear models (GLM) using both $\cos(\phi_x)$ and $\sin(\phi_x)$ by Penny *et al.* [Penny et al., 2008]. Other approaches employed a measure of coherence [Colgin et al., 2009] or the phase-locking value [Lachaux et al., 1999]. All these last three approaches offer metrics which are independent of the phase at which the maximum amplitude occurs. The methods of Tort *et al.* [Tort et al., 2010], Ozkurt *et al.* [Özkurt and Schnitzler, 2011], and Penny *et al.* [Penny et al., 2008] will be considered for comparison in our experiments.

As one can see, there is a long list of methods to quantify CFC in neural time series. Yet, a number of limitations which can significantly affect the outcomes and interpretations of neuroscientific findings exist with these approaches. For example, in typical PAC analysis, a systematic bias rises where one constructs the so-called *comodulogram*. A comodulogram is obtained by evaluating the chosen metric over a grid of frequency f_x and f_y. This bias emerges from the choice of the bandpass filter, which involves the critical choice of the bandwidth Δf_y. It has been reported several times that to observe any amplitude modulation, the bandwidth of the fast oscillation Δf_y has to be at least twice as high as the frequency of the slow oscillations f_x: $\Delta f_y > 2f_x$ [Berman et al., 2012, Dvorak and Fenton, 2014]. As a comodulogram uses different values for f_y, many studies have used a variable bandwidth, by taking a fixed number of cycles in the filters. The bandwidth is thus proportional to the center frequency: $\Delta f_y \propto f_y$. This choice leads to a systematic bias, as it hides any possible coupling below the diagonal $f_y = 2f_x/\alpha$, where $\alpha = \Delta f_y / f_y$ is the proportionality factor. Other studies have used a constant bandwidth Δf_y; yet this also biases the results towards the low driver frequency f_x, considering that it hides any coupling with $f_x > \Delta f_y/2$. A proper way to build a comodulogram would be to take a variable bandwidth $\Delta f_y \propto f_x$, with $\Delta f_y > 2f_x$. However, this is not common practice as it is computationally very demanding, because it implies to bandpass filter y again for each value of f_x.

Another common issue arises with the use of the Hilbert transform to estimate the amplitude and the phase of real-valued signals. Such estimations rely on the hypothesis that the signals x and y are narrow-band, i.e. almost sinusoidal. However, numerous studies have used this technique on very wide-band signals such as the entire gamma band (80-150 Hz) [Canolty et al., 2006] (see other examples in [Chavez et al., 2006]). The narrow-band assumption is debatable for high frequency activity and, consequently, using the Hilbert transform may yield non-meaningful amplitude estimations, and potentially poor estimations of PAC [Chavez et al., 2006, Dvorak and Fenton, 2014]. Note also that, in this context, wavelet-based filtering is equivalent to the Hilbert transform [Quiroga et al., 2002, Bruns, 2004], and therefore does not provide a more valid alternative option.

Besides these issues of filtering and inappropriate use of Hilbert transforms, Hyafil [Hyafil, 2015] also warned that certain choices of bandwidth Δf_y might mistake phase-frequency coupling for PAC, or create spurious amplitude-amplitude coupling; see also the more recent work in [Aru et al., 2015] for discussion and more practical recommendations for PAC analysis.

Here we advocate that the DAR models detailed in the next sections address a number of the limitations just mentioned. They do *not* use bandpass filter or Hilbert transform on the high frequencies y. They introduce a measure of *goodness of fit*, through the use of a probabilistic signal model whose quality can be assessed by evaluating the *likelihood* of the data under the model. In practice, the likelihood quantifies how much variance of the signal can be explained by the model, and is similar to the R^2 coefficient in generalized linear models (GLM). To the best of our knowledge, the only related model-based approach to measure PAC used GLM [Penny et al., 2008]. With GLM, however, the modeling part is done independently on each signal y_f, which is the band-pass filtered version of y around frequency f. For each of these frequencies f a different model is fitted. By doing so, a GLM approach cannot model the wide-band signal y as it is limited to multiple estimations frequency by frequency bin. This largely limits the use of the likelihood to compare models or parameters. On the contrary, we propose to model y globally, without filtering it in different frequency bands.

To conclude this section, and to position this work in the broader context of modeling approaches for neuroscience data, we would like to stress that our proposed method can be considered as an encoding model for CFC, as opposed to a decoding model [Kay et al., 2008, Naselaris et al., 2011, Huth et al., 2016]. Indeed, our model reports how much empirical data can be explained and by doing so enables us to test neuroscience hypotheses in a principled manner [Naselaris et al., 2011].

Non-linear AR models beyond Neuroscience

The literature on the use of non-linear auto-regressive (AR) models is quite large and covers fields such as audio signal processing and econometrics. For instance, AR models with conditional heteroskedasticity (ARCH [Engle, 1982], GARCH [Bollerslev, 1986]) are extremely popular in econometrics where they are used to model signals whose overall amplitude varies as a function of time. Here, however, in the context of CFC and PAC, one would like to model variations in the spectrum itself, such as shifts in peak frequencies (a.k.a. frequency modulations) or changes in amplitude only within certain frequency bands (a.k.a. amplitude modulations). To achieve this, one idea is to define a linear AR model, whose coefficients are a function of time and change slowly depending on a *non-linear* function of the signal.

The first models based on this idea are SETAR models [Tong and Lim, 1980], which switch between several AR models depending on the amplitude of the signal with respect to some thresholds. To get a smoother transition between regimes, SETAR models have inspired other models like EXPAR [Haggan and Ozaki, 1981] or STAR [Chan and Tong, 1986], in which the AR coefficients change continuously depending on a non-linear function of the past of the signal.

These models share the same underlying motivation as the DAR models described below but, crucially, DAR models can be designed and parametrized to capture PAC phenomena independently of the phase in the driving signal at which the high frequency content is the strongest. In other words, DAR models can work equivalently well if the high frequency peaks are in the troughs, the rising phase, the decreasing phase or the peaks of the low frequency driving signal. Moreover, as DAR models do not require to infer the driving behavior from the signal itself and rather rely on the prior knowledge of the slow oscillation, the inference is significantly faster and more robust.

Methods

In this section, we first define our statistical signal model, explain how we estimate its parameters on a signal y and its driver x, and demonstrate how one can infer hyper-parameters by comparing the likelihood of several models. Then, we detail how to use this model on neurophysiological signals by presenting the preprocessing steps and showing how to make the model invariant to the phase of the coupling. We also detail how to express the power spectral density (PSD) conditionally to the driver's phase, which allows for a fine-grained analysis of the signal's spectral properties and to build comodulograms. Finally, we present our protocol to simulate signal with CFC and the empirical datasets used to validate our method.

DAR models - General presentation

Model definition

An auto-regressive (AR) model specifies that a signal y depends linearly on its own p past values, where p is the *order* of the model:

$$\forall t \in [p+1, T] \qquad y(t) + \sum_{i=1}^{p} a_i y(t-i) = \varepsilon(t) \ . \tag{1}$$

Here, T is the length of the signal and ε is the *innovation* (or *residual*) modeled with a Gaussian white noise: $\varepsilon(t) \sim \mathcal{N}(0, \sigma(t)^2)$. To extend this AR model to a non-linear model, one can assume that the AR coefficients a_i are non-linear functions of a given exogenous signal x, here called the *driver*:

$$\forall i \in [1, p] \qquad a_i(t) = g_i(x(t)) \ . \tag{2}$$

What was proposed in [Grenier, 2013, Grenier, 1983] is to consider that the non-linear functions g_i are polynomials:

$$\forall t \in [p+1, T] \qquad g_i(x(t)) = \sum_{k=0}^{m} a_{ik} x(t)^k \ , \tag{3}$$

where $x(t)^k$ is $x(t)$ to the power k. Note that one could imagine more complex models for the g_i, yet empirical evidence confirms that polynomials are rich enough for the different neurophysiological signals considered in this work (see S1 Fig).

Inserting the time-dependent AR coefficients (2) into the AR model (1), we obtain the following equation:

$$\forall t \in [p+1, T] \qquad y(t) + \sum_{i=1}^{p} \sum_{k=0}^{m} a_{ik} x(t)^k y(t-i) = \varepsilon(t) \ . \tag{4}$$

This can be simplified and rewritten as:

$$\forall t \in [p+1, T] \qquad y(t) + A^\top \tilde{Y}(t) = \varepsilon(t) \ , \tag{5}$$

where $A \in \mathbb{R}^{p(m+1)}$ is a one-dimensional vector composed of the scalars a_{ik} and $\tilde{Y}(t) \in \mathbb{R}^{p(m+1)}$ is a one-dimensional vector composed of the regressors $x(t)^k y(t-i)$. We note $\hat{A}$ the estimated value of A and $A^\top$ its transpose.

We also consider a time-varying innovation variance $\sigma(t)^2$ driven by x. This corresponds to the assumption that the power of the signal at time t depends on the driver at this same instant. Since the standard deviation is necessarily positive, we use the following polynomial model for its logarithm:

$$\forall t \in [p+1, T] \qquad \log(\sigma(t)) = \sum_{k=0}^{m} b_k x(t)^k = B^\top X(t) \ , \tag{6}$$

where $B \in \mathbb{R}^{m+1}$ is a one-dimensional vector composed of the scalars b_k.

This model is called a driven auto-regressive (DAR) model [Dupré la Tour et al., 2017]. DAR models are thus a natural extension of well-known AR models and enable analyzing non-stationary or non-linear spectral phenomena observed in neurophysiological time-series. One methodological concern is that the AR coefficients cannot be arbitrary as they should define stable models (see [Proakis and Manolakis, 1996], section 3.6). In our case there is an extra difficulty because they are changing over time. To guarantee that each set of coefficients defines a stable filter, it is possible to re-parametrize the model. This requires significantly more complex derivations and leads to a slightly slower optimization, while providing equivalent empirical observations. For simplicity, this manuscript will focus on the simplest model dedicated to CFC analysis, and we refer to [Dupré la Tour et al., 2017, Grenier and Omnes-Chevalier, 1988] for readers interested in more details.

Model estimation

DAR model equation (4) is non-linear with respect to the given signals x and y, yet it is linear with respect to the regressors $x(t)^k y(t-i)$. Therefore after computing the regressors, it is possible to obtain an analytical expression of the parameters in A.

As the innovation $\varepsilon(t)$ is assumed to be a Gaussian white noise, the model likelihood L is obtained via:

$$L = \prod_{t=p+1}^{T} \frac{1}{\sqrt{2\pi\sigma(t)^2}} \exp\left(-\frac{\varepsilon(t)^2}{2\sigma(t)^2}\right) \tag{7}$$

or

$$-2\log(L) = T\log(2\pi) + \sum_{t=p+1}^{T} \frac{\varepsilon(t)^2}{\sigma(t)^2} + 2\sum_{t=p+1}^{T} \log(\sigma(t)) \ .$$

DAR models are estimated by likelihood maximization. Here, if the innovation variance $\sigma(t)^2$ is considered fixed, maximizing this function boils down to minimizing the sum of squares of the $\varepsilon(t)$. Since $\varepsilon(t)$ is the residual, the inference of the parameters amounts to maximizing the variance explained by the model. To start the maximization, we first assume that the innovation variance is fixed and equal to the signal's empirical variance. The likelihood then reads:

$$-2\log(L) = C + \sum_{t=p+1}^{T} \frac{\varepsilon(t)^2}{\sigma(t)^2} = C + \sum_{t=p+1}^{T} \frac{1}{\sigma(t)^2}\left(y(t) + A^\top \tilde{Y}(t)\right)^2 \ , \tag{8}$$

where C is a constant. We can thus estimate the DAR coefficients $A \in \mathbb{R}^{p(m+1)}$ by solving the following linear system (a.k.a. normal equations):

$$\left(\sum_{t=p+1}^{T} \frac{1}{\sigma(t)^2} \tilde{Y}(t)\tilde{Y}(t)^\top\right) \hat{A} = -\left(\sum_{t=p+1}^{T} \frac{1}{\sigma(t)^2} \tilde{Y}(t) y(t)\right) \ . \tag{9}$$

Given $\hat{A}$, one can then estimate the vector $B \in \mathbb{R}^{m+1}$ from the residual $\varepsilon(t)$, maximizing the likelihood using off-the-shelf optimization techniques adapted to differentiable problems. We use a Newton-Raphson procedure (see more details in [Grenier, 2013]). The instantaneous variance is then updated with B using (6). One can then iterate between the estimation of the coefficients A and B. In our experiments, we observed that one or two iterations are sufficient.

Model selection

DAR models are probabilistic, which offers a significant advantage over traditional PAC metrics. Indeed, one can evaluate the likelihood of the model L (7), which quantifies the *goodness of fit* of the data given the model, and select the hyper-parameters that maximize it. To compare models with different number of hyper-parameters (i.e. degrees of freedom), different information criteria can be used such as the Akaike Information Criterion (AIC) [Akaike, 1998] or the Bayesian Information Criterion (BIC) [Schwarz et al., 1978], which read respectively $\mathrm{AIC} = -2\log(L) + 2d$ and $\mathrm{BIC} = -2\log(L) + d\log(T)$. The integer d is the number of degrees of freedom of the model, which for DAR models equals to $d = (p+1)(m+1)$. The BIC has been used extensively for hyper-parameter selection in AR models, and enables the joint estimation of both p and m (see S2 Fig for simulation results).

As the statistical inference with a DAR model is very fast, one can easily estimate models on a full grid of hyper-parameters (p and m) and keep the model that leads to the lowest BIC. Another alternative used in our experiments is the use of cross-validation (CV), consisting of splitting the recorded signals into two sets: a training signal to estimate parameters, and a test signal to evaluate the likelihood of new data. In practice, CV requires that a large amount of data be left for model testing, and that the two datasets be separated by a minimum delay to ensure data independence [Arlot et al., 2010]. In S3 Fig, we present model selection based on cross-validation, applied on real signals.

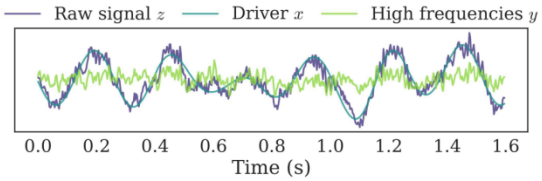


Fig 1. Example of a preprocessed signal. The driver x is extracted from the raw signal z with a bandpass filter, and then subtracted from z to give y. A PAC effect is present as we see stronger high frequency oscillations at the peaks of the driver. The signal y is presented here as it looks before the temporal whitening step.

DAR models for CFC estimations on brain data

Data preprocessing: Extracting the driver

In this section, we describe the preprocessing applied to the raw signal z to obtain the signal y and the driver x.

First, the raw signal was downsampled to an appropriate sampling frequency f_s (we used 333 Hz and 625 Hz in our examples), as we only consider frequencies up to $f_s/2$. Power line noise (50 or 60 Hz for French and US datasets, respectively) and its harmonics were removed. As power line noise fluctuates over time, the frequency of the power line noise was estimated by projecting the signal onto sinusoids, and the maximum-power projection was subtracted from the raw signal. When necessary, a high-pass filter was used to detrend the signal (typically at 1 Hz) [Bigdely-Shamlo et al., 2015].

A bandpass filter was then used to extract the driver x. The center frequency f_x and bandwidth Δf_x for this filter will be discussed in details in the Results section. The filter equation is:

$$w(t) = b(t)\cos(2\pi f_x t) \ , \tag{10}$$

where $b(t)$ is a Blackman window of order $\lfloor 1.65 * f_s/\Delta f_x \rfloor * 2 + 1$, chosen to have a bandwidth of Δf_x at -3 dB (the filter attenuation is 50% at $f_x \pm \Delta f_x/2$). The filter is zero-phase since it is symmetric.

Then, the driver x is subtracted from the raw signal z to create the signal $y = z - x$. Note that the signal y now contains a frequency gap around f_x, which can be a nuisance for the AR estimate that provides a compact model for the broad band power spectrum density of the signal. To solve this issue, we fill this gap by adding a Gaussian white noise filtered with the filter w, and adjusted in energy to have a smooth power spectral density (PSD) in y. Such a preprocessing step is also commonly used when working with multivariate auto-regressive models (MVAR) on neurophysiological signals corrected for power line noise.

Finally, we whiten y with a linear AR model, by applying the inverse AR filter to the signal. This temporal whitening step is not necessary, yet it reduces the need for high order p in DAR and therefore reduces both the computational cost and the variance of the model. After this whitening step, the PSD of y is mostly flat, and DAR models only contain the modulation of the PSD. Fig. 1 shows a time sample of a preprocessed signal, taken from a rodent striatal LFP signal (see below for a description of the data). One can see how the slowing varying driving signal follows the original raw signal and how the high frequencies bursting on the peak of the slow oscillation remain in the processed signal y. The general processing pipeline is summarized in Fig. 2(a).

Model selection: Optimizing the filtering parameters

During the preprocessing, the driver x_c is extracted from the preprocessed signal z through a band-pass filter, with center frequency f_x and bandwidth Δf_x. To select automatically these two parameters, we perform a grid-search and select the driver which yields the highest likelihood. Importantly, the highest likelihood does

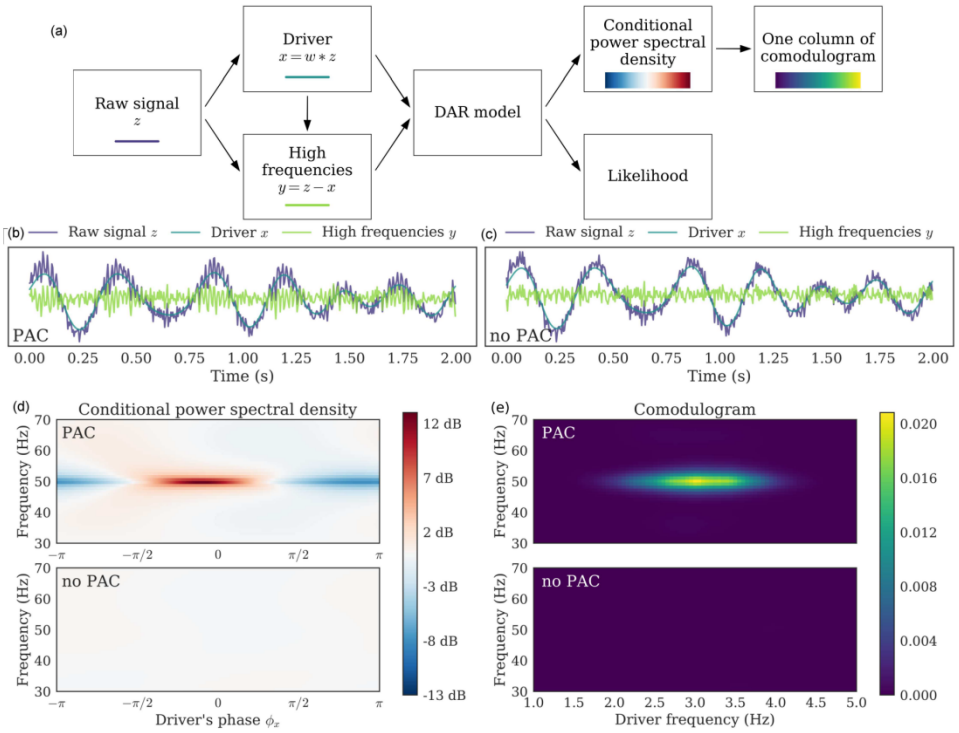


Fig 2. Example on simulated signals: Pipeline, signal, conditional PSD, and comodulogram. (a) Pipeline of the method. We applied it with $(p, m) = (10, 1)$ on two simulated signals: (b) Simulated signal with PAC and (c) Simulated signal without PAC. (d) From a fitted model, we computed the PSD conditionally to the driver's phase ϕ_x; each line is centered independently to show amplitude modulation. PAC can be identified in the fluctuation of the PSD as the driver's phase varies: around 50 Hz, the PSD has more power for one phase than for another. This figure corresponds to a single driver's frequency $f_x = 3.0$ Hz. (e) Applying this method to a range of driver's frequency, we build a comodulogram, which quantifies the PAC between each pair of frequencies.

not necessarily correspond to the highest PAC score, but rather to the model with the highest goodness of fit or explained variance. This detail is crucial, as optimizing for the best fit is a legitimate data-driven approach, whereas optimizing for the highest PAC score is statistically more questionable. One additional benefit of our approach is that other types of filters can be compared using a common methodological approach.

For a fair comparison of the drivers, we need to evaluate the model fitting on the exact same signal y. To do so, we remove from y all the possible drivers tested on this grid-search using a high-pass filter above maximum center frequency f_x. Once again, we fill the frequency gap by adding a Gaussian white noise filtered with a low-pass filter complementary to the high-pass filter.

Phase invariance: Removing the bias using a complex driver

We consider PAC where high frequencies are coupled with a low frequency driving signal. We denote by ϕ_0 the preferred phase of the coupling, i.e. the phase of the driver that corresponds to the maximum amplitude of the high frequencies. When $\phi_0 = 0$, the high frequency bursts happen in the peaks of the driver, whereas when $\phi_0 = \pi$, they happen in the troughs.

The driver extracted as described in previous subsection is real-valued, and the filter used to extract it is based on a cosine. As the cosine has the same value for ϕ and $\pi - \phi$, the PAC estimation is biased and underestimated when ϕ_0 is not equal to 0 or π. Indeed, the model mixes the contribution of ϕ and $\pi - \phi$, which attenuates the effect. This bias was already reported in PAC estimation by [Bruns and Eckhorn, 2004], and the solution proposed by [Penny et al., 2008] was to use both cosine and sine to have PAC estimates that are invariant to the preferred phase ϕ_0 of the low frequency oscillation.

The same technique can be used on the DAR models. We not only filter the raw signal with $w(t) = b(t)\cos(2\pi f_x t)$ to obtain x, but also with $\bar{w}(t) = b(t)\sin(2\pi f_x t)$ to obtain $\bar{x}$, creating a complex-valued driver $x_c = x + j\bar{x}$ where j denotes the complex number, $j^2 = -1$. With a complex-valued driver, DAR models are naturally extended by adding more regressors:

$$\forall i \in [1, p] \qquad a_i(t) = \sum_{0 \leq k+l \leq m} a_{ikl}\, x(t)^k \bar{x}(t)^l \ . \tag{11}$$

Instead of $p(m+1)$ regressors, we now have $p(m+1)(m+2)/2$ regressors, and the number of degrees of freedom of the model is now $d = (p+1)(m+1)(m+2)/2$. Typical values for m are below 3, so the number of parameters stays within a reasonable range despite the squared dependence in m. The number of parameters in the model remains much lower than the number of time samples and the estimation problem stays therefore well-posed. The innovation variance model is also updated accordingly:

$$\log(\sigma(t)) = \sum_{0 \leq k+l \leq m} b_{kl}\, x(t)^k \bar{x}(t)^l \tag{12}$$

A validation result is presented in the supporting information (S4 Fig), showing that indeed it removes the preferred-phase bias.

Conditional spectrum: Deriving the PSD from the model

After preprocessing and model estimation, we can compute the *conditional* power spectral density (PSD) of the model, noted $\mathrm{PSD}_y(x_{c0})$ since it is here a function of the driver's value x_{c0}. For a given complex driver's value x_{c0}, we compute the AR coefficients $a_i(x_{c0})$ using Eq. (11), along with the innovation's standard deviation $\sigma(x_{c0})$ using Eq. (12). Since AR models with time-varying coefficients are locally stationary [Dahlhaus, 1996], we can compute the PSD at a frequency f with:

$$\mathrm{PSD}_y(x_{c0})(f) = \sigma(x_{c0})^2 \left| \sum_{i=0}^{p} a_i(x_{c0}) e^{-j2\pi f i} \right|^{-2} \tag{13}$$

where $j^2 = -1$ and $a_0(x_0) = 1$. We thus have a different PSD for each driver value x_c, i.e. for each time instant, as the driver $x_c(t)$ fluctuates in time. Note that the PSD is a function of the driver value and not really its phase, unless the driver has been normalized in amplitude. The varying AR coefficients $a_i(t)$ parametrize the fluctuations of the shape of the PSD, whereas the varying innovation's standard deviation $\sigma(t)$ is responsible for the absolute fluctuations in power over all frequencies.

In Fig. 2(b and c), one can see two simulated signals: one with PAC on the left (b) and one without PAC on the right (c). The simulation process, which does not make use of DAR model, will be detailed and discussed in a following section. In Fig. 2(d), one can see an example of $\mathrm{PSD}_y(x_c)$ as a function of the driver's value $x_c = \rho\exp(j\phi)$. In practice we set ρ to the median of the driver's amplitude, $\rho = \mathrm{median}(|x_c|)$, and ϕ spans the entire interval $[-\pi, \pi]$. The PAC is identified by the modulation of the spectrum with respect

to the driver's phase. Note that with DAR models, the amplitude modulation is modeled jointly on the entire spectrum, as opposed to frequency band by frequency band in most other PAC metrics.

Using the fast Fourier transform (FFT) in Eq. (13), we are able to compute the PSD for a broad range of frequencies $f \in [0, f_s/2]$. As model estimation is also very fast, we obtain the conditional PSD very quickly: With $(p, m) = (10, 1)$ on a signal with $T = 10^5$ time points, the entire DAR estimation and PSD computation lasts 0.12 s.

Comodulogram: Quantifying the modulation at each frequency

We now detail how comodulograms can be derived from DAR models. To quantify the coupling at a given frequency f, we compute $\text{PSD}_y(x_c)(f)$ for a range of n artificial driver's values $x_c(k)$ located on a circle of radius $\rho = \text{median}(|x_c(t)|)$, and we normalize it to sum to one:

$$\forall k \in [1, n] \qquad p_f(k) = \frac{\text{PSD}_y(\rho e^{j2\pi k/n})(f)}{\sum_{k=1}^{n} \text{PSD}_y(\rho e^{j2\pi k/n})(f)} \quad . \tag{14}$$

Then, we use the same method as [Tort et al., 2010] to measure the fluctuation of $p_f(k)$: we compute the Kullback-Leibler divergence between $p_f(k)$ and the uniform distribution $q(k) = 1/n$, and we normalize it with the maximum entropy $\log(n)$:

$$M(f) = \frac{1}{\log(n)} \sum_{k=1}^{n} p_f(k) \log\left(\frac{p_f(k)}{q(k)}\right) \quad . \tag{15}$$

This metric M is between 0 and 1 and allows us to quantify the non-flatness of $p_f(k)$, i.e. the amplitude modulation, frequency by frequency. We can then build a *comodulogram*, a figure commonly used in the literature which depicts the PAC metric for a grid of frequency (f_x, f). When derived from DAR models as described here, the comodulogram is not affected by the systematic biases presented in the introduction.

Fig. 2(e) shows the comodulogram of the two simulated signals, one with and one without PAC. As expected the DAR model reports no coupling when there is none, and reports here a strong coupling between 3 and 50 Hz when it is actually present.

Directionality: Estimating a delay between the signal and the driver

The DAR model in Eq. (4) uses the driver at time t to parametrize the AR coefficients with the assumption that, in PAC, high frequency activity is modulated by the driving signal with no delay. However, one could legitimately ask whether slow fluctuations in oscillatory activity precedes the amplitude modulation of the fast activity or, conversely, whether high frequency activity precedes low frequency fluctuations. This is what we will refer to as *directionality estimation*.

To assess the problem of directionality estimation with DAR models, we capitalize on the likelihood function. After observing some PAC between y and x, we introduce a delay τ in the driver so as to estimate the goodness of fit of a DAR model between y and x_τ, where $x_\tau(t) = x(t - \tau)$. Using a grid of values for the delay, one can report the value τ_0 that corresponds to the maximum likelihood. In practice, as we use non-causal filters to extract the driver, and because the AR model uses samples from the past of the y, one might wonder if there is any positive or negative bias in our delay estimation. To address this issue, we apply our analysis on both forward and time-reversed signals, and sum the two models log-likelihood. In this way, the filtering bias is strongly attenuated because it similarly affects both forward and time-reversed models. If the best delay is positive, it means that the past driver yields a better fit than the present driver, i.e. that the slow oscillation precedes the amplitude modulation of the fast oscillation. Inversely, a negative delay means that the amplitude modulation of the fast oscillation follows changes in the slow oscillation.

It is worth noting that the estimation of the delay τ_0 shall not be considered as an alternative way to estimate the preferred phase ϕ_0 defined in a previous section. Although the preferred phase and the delay would be identical should the driver be a perfect stationary sinusoid, in non-stationary neural systems in which the driver's instantaneous frequency may fluctuate with time, the preferred phase and delay will likely

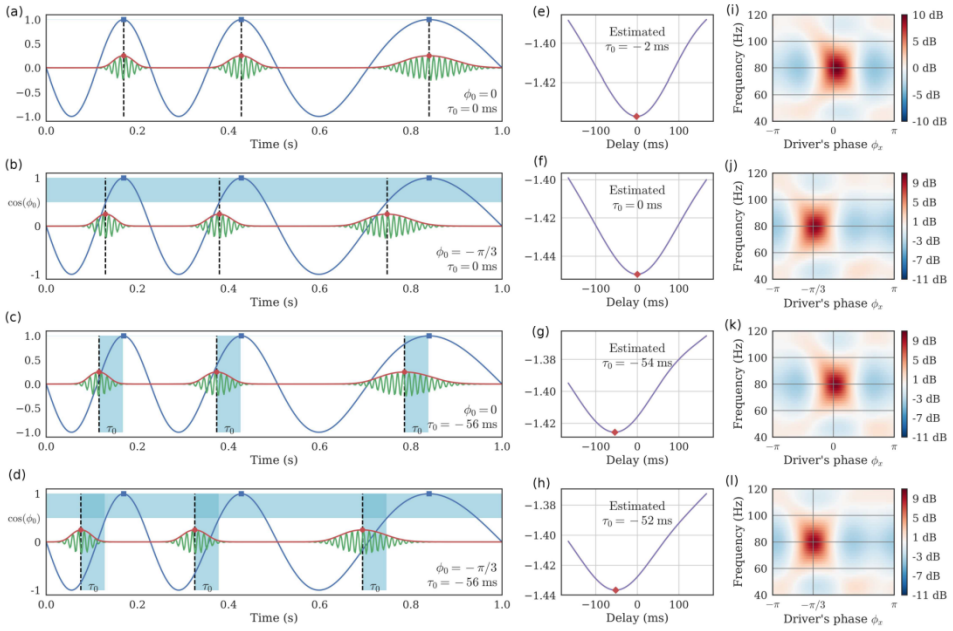


Fig 3. Preferred phase and temporal delay are different. The temporal delay τ is distinct from the preferred phase ϕ_0. (a) When both are equal to zero, the high frequency bursts happen in the driver's peaks. (b) When $\tau = 0$ and $\phi \neq 0$, the bursts are shifted in time with respect to the driver's peaks, and this shift *varies* depending on the instantaneous frequency of the driver. (c) When $\tau \neq 0$ and $\phi = 0$, the bursts are shifted in time with respect to the driver's peaks, and this shift is *constant* over the signal. In this case, note how the driver's phase corresponding to the bursts varies depending on the instantaneous frequency of the driver. (d) τ and ϕ_0 can also be both non-zero. (e-h) Negative log-likelihood of DAR models, fitted with different delays between the driver and the high frequencies. The method correctly estimates the delay even when $\phi_0 \neq 0$. (i-l) PSD conditional to the driver's phase, estimated through a DAR model with the best estimated delay. The maximum amplitude occurs at the phase ϕ_0.

differ. Such a scenario can occur either because signal waveforms are not perfect sinusoids [Cole et al., 2016] or because the driver actually changes [Tort et al., 2008]. In other words, τ_0 is a time delay which is identical at all time and corresponds to different phase shifts $\phi(t) = \tau_0 * (2\pi f(t))$ which depend on the instantaneous frequency $f(t)$. On the contrary, ϕ_0 is a phase shift that is constant over time and which corresponds to different time delays $\tau(t) = \phi_0/(2\pi f(t))$ which also depend on the instantaneous frequency.

Fig. 3 illustrates these specific points and disentangles the two distinct notions.

Like most PAC metrics, DAR models are invariant with respect to the preferred phase. Both the strength of the coupling and the model likelihood are unchanged with respect to ϕ_0. On the contrary, all PAC metrics including DAR models are strongly affected by a time delay τ when the driver is not a perfect sinusoid. This attenuates the coupling and can potentially modify artificially the preferred phase. When the delay is too large, all metrics would measure zero coupling. This is in fact what justifies the use of surrogate techniques that introduce a large time shift to quantify the variance of the measure in the absence of coupling [Canolty et al., 2006, Tort et al., 2010, Aru et al., 2015]. Fig. 3(e-h) provides the delay estimated with DAR models, obtained by maximizing the likelihood over a grid of delays, as explained previously. In Fig. 3(i-l), we show

the conditional PSD of the model obtained with the best estimated delay as well as the estimated preferred phase ϕ_0. Using the likelihood, our method is thus able to estimate this delay, thus improving PAC estimation and extending PAC analysis.

The question of delay estimation has been previously addressed in the literature with measures borrowed from information theory [Besserve et al., 2010]. For instance, Transfer entropy (TE) [Schreiber, 2000, Wibral et al., 2013, Park et al., 2015] was adapted to PAC by directly computing the envelop of both the slow and the fast oscillations. When variables follow a Gaussian distribution, TE is equivalent to Granger causality [Barnett et al., 2009], but Granger causality can fail when the signal-to-noise ratios of the two signals are different [Nolte et al., 2010], which is often the case for the driver and the envelop of the high frequencies. Hence, it remains to be investigated how TE compares with DAR models in terms of performance. It has also been addressed with another method called Cross-Frequency Directionality (CFD) [Jiang et al., 2015], which is based on the phase slope index (PSI) [Nolte et al., 2008]. CFD is not designed to measure a delay; thus, we modified it so as to compare it *qualitatively* to our approach (see Results). If such delay estimation results may not reflect pure causality [Aru et al., 2015], for instance because of the transitivity of correlation, they nevertheless improve the analysis of PAC going one step further by estimating the delay between the coupled components.

Data

Simulated signals

In our experiments, methods comparisons and model evaluations, we make an extensive use of simulated signals showing some PAC. The simulations are generated as follows: we simulate a coupling between $f_x = 3\,\text{Hz}$ and $f_y = 50\,\text{Hz}$, with a sampling frequency $f_s = 240\,\text{Hz}$, during T time points (T depends on the experiment). We do not use a perfect sinusoidal driver, as using such an ideal oscillatory signal is over-simplistic. Drivers can indeed have a larger band as suggested next by our experiments. Finally strong empirical evidence suggests that neurophysiological signals have complex morphologies that can be overlooked when studied as ideal sinusoids [Cole and Voytek, 2017].

To construct a time-varying driver peaking at 3 Hz, we bandpass filter a Gaussian white noise at a center frequency $f_x = 3\,\text{Hz}$ and with a bandwidth $\Delta f_x = 1\,\text{Hz}$, using the same filter detailed in the preprocessing section. The smaller the bandwidth, the closer the driver is to a perfect sinusoid. We normalize the driver to have unit standard deviation $\sigma_x = 1$. We then modulate the amplitude a_y of a sinusoid at, for example, $f_y = 50\,\text{Hz}$, using a sigmoid on the driver x:

$$a_y(t) = \frac{1}{1 + \exp(-\lambda x(t))} \ , \tag{16}$$

with a sharpness set to $\lambda = 3$. By doing so, the amplitude varies between 0 and 1 depending on the driver's value. We normalize the signal y to have a standard deviation $\sigma_y = 0.4$. We then add to the signal y both the driver x and a Gaussian white noise with a standard deviation $\sigma_\varepsilon = 1$. Note that this simulation procedure is not based on a DAR model. In other words, we do not validate our model using signals that fall perfectly into the category of stochastic signals that are synthesized with a DAR model.

Empirical datasets

We tested DAR models on three different datasets, with invasive recordings in humans and rodents:

1. A rodent local field potential (LFP) recording in the dorso-medial striatum from [Dallérac et al., 2017](supplementary figure 2), (1800 seconds down-sampled at 333 Hz)
2. A rodent LFP recording in the hippocampus from [Khodagholy et al., 2015] collected during rapid eye movement (REM) sleep (100 seconds down-sampled at 625 Hz)
3. A human electro-corticogram (ECoG) channel in auditory cortex from [Canolty et al., 2006] (730 seconds down-sampled at 333 Hz)

We refer the reader to corresponding articles for more details on the recording modalities of these neurophysiological signals.

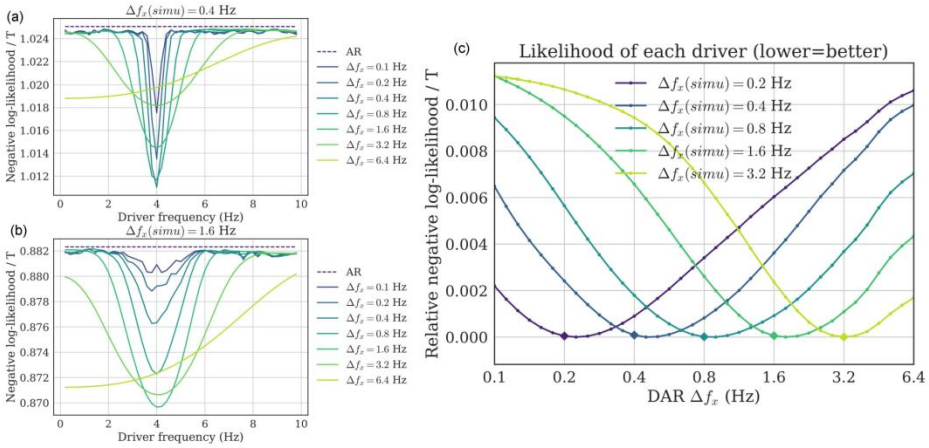


Fig 4. Driver's frequency and bandwidth selection (simulations). (a,b) Two examples of grid-search over both f_x and Δf_x, over simulated signals with (a) $\Delta f_x(simu) = 0.4\,\text{Hz}$ and (b) $\Delta f_x(simu) = 1.6\,\text{Hz}$. For all bandwidths (except 6.4 Hz), the negative log-likelihood is minimal at the correct frequency $f_x = 4\,\text{Hz}$. (c) To see more precisely the bandwidth estimation, we plot the negative log-likelihood (relative to each minimum for readability), for several ground-truth bandwidth $\Delta f_x(simu)$, with $f_x = 4\,\text{Hz}$. For each line, the minimum correctly estimates the ground-truth bandwidth (depicted as a diamond), showing empirically that the parameter selection method gives satisfying results.

Results

We now illustrate the benefits of DAR models using simulated data and neurophysiological recordings in rats and humans. Our primary goal is to demonstrate that assessing the validity of the model through its likelihood is a statistically principled approach that allows to reveal new insights on empirical data with interesting neuroscientific consequences.

Model selection and data-driven estimation of the driving signal

In this section, we present the outcome of using the model selection procedure to estimate the best filtering frequency f_x and bandwidth Δf_x to extract the driver x. We first describe the outcome on simulated signals (ground truth) and then on empirical datasets.

Simulation study: Validating the approach

To validate our approach, we tested on simulated signals whether both center frequency f_x and bandwidth Δf_x were correctly estimated. For this, we simulated signals using $f_x(simu) = 4\,\text{Hz}$ and $\Delta f_x(simu) \in [0.2, 0.4, 0.8, 1.6, 3.2]\,\text{Hz}$. To mimic the duration of our three real signals, we used $T = \lfloor 100 f_s \rfloor$ time points, for a duration of 100 seconds.

As the noise mainly affected the amplitude modulation but barely altered the driver, we also added a Gaussian white noise low-pass filtered at 20 Hz, and scaled to have a PSD difference of 10 dB at $f_x = 4\,\text{Hz}$ between the driver and the noise. In this way, the driver was also altered by noise.

Given the simulated signal, we performed a grid-search over f_x and Δf_x, extracting the driver, fitting the model and computing the likelihood of the model. Results are presented in Fig. 4. As expected, the center

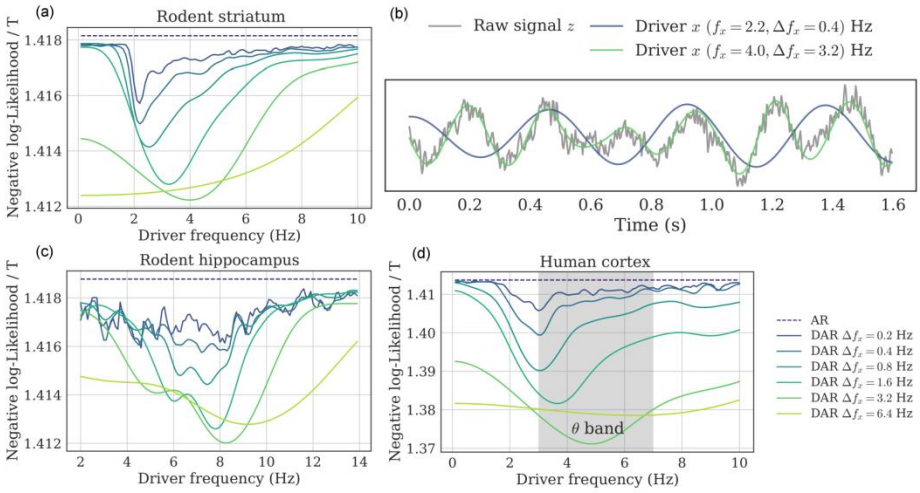


Fig 5. Driver's frequency and bandwidth selection (real signals). (a,c,d) Negative log-likelihood of the fitted model for a grid of filtering parameters f_x and Δf_x: (a) rodent striatum, (c) rodent hippocampus, (d) human auditory cortex. The optimal bandwidth was very large (3.2 Hz), and the optimal center frequency changed as the bandwidth increased, suggesting that the optimal driver had a wide asymmetrical spectrum. (b) This portion of the rodent striatal signal shows two examples of driver with different bandwidths: The wide-band driver better follows the raw signal, and *independently* also leads to a better fit in DAR models.

frequency f_x was correctly estimated for all bandwidths Δf_x. More importantly, the negative log-likelihood was minimal at the correct simulated bandwidth.

This simulation confirms that the likelihood can be used to estimate the correct parameters for the driver's filtering step, and that does not present any obvious bias in the estimation.

Driver estimates on human and rodent recordings

The outcome of the model selection procedure on the three real neurophysiological signals are reported in Fig. 5. Two general observations can be made. First, for all three signals, the optimal bandwidth was relatively large (3.2 Hz). Second, the optimal center frequency changed as we increased the bandwidth. Interestingly, this phenomenon was not observed in the simulation study, suggesting that in real data the optimal driver is wide-band and has an asymmetrical spectrum. In other words, the driver's frequency is not precisely defined, and a large band-pass filter should be preferred to extract the driver.

These observations thus question which parameters should be used in practice. The classical approach is to build a comodulogram to select the best driver frequency f_x, choosing arbitrarily the bandwidth Δf_x. This bandwidth is typically quite narrow, e.g. 0.4 Hz, and chosen to clearly isolate the maximum frequency in the comodulogram. However, this approach relies on the assumption that the driver is nearly sinusoidal, which is unrealistic [Cole and Voytek, 2017]. On the contrary, our data-driven approach selects the frequency and bandwidth that lead to the highest goodness of fit of our model.

To further describe the implications of our observations, we focused on the rodent striatal LFP signal. We compared an arbitrary choice of driver bandwidth (0.4 Hz) and the optimal choice with respect to the model likelihood (3.2 Hz). For these two bandwidths, we selected the optimal center frequencies, 2.2 Hz and 4.0 Hz respectively. Fig. 5b shows a time sample of the raw signal and the two extracted drivers. The wide-band driver (in green) followed very well the raw signal, which was not a perfect sinusoid. On the contrary, the

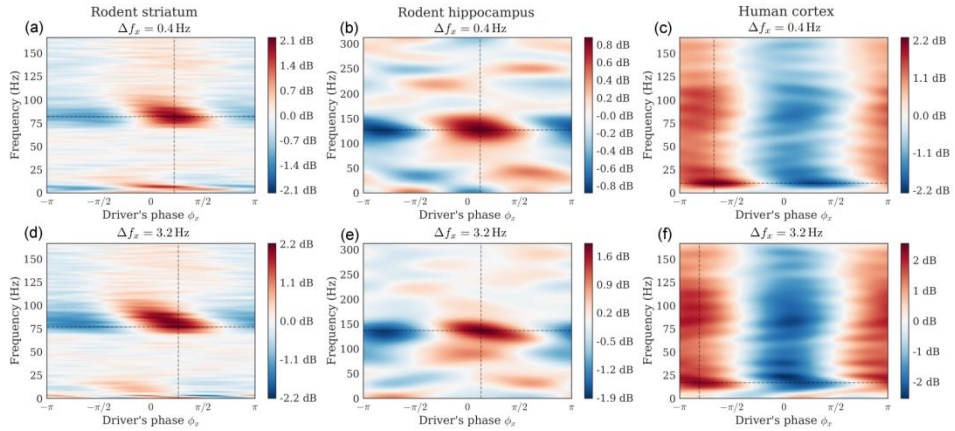


Fig 6. PSD conditional to the driver's phase. Dataset: Rodent striatum (a,d), rodent hippocampus (b,f), human auditory cortex (c,g). We derive the conditional PSD from the fitted DAR models. In one plot, each line shows at a given frequency the amplitude modulation with respect to the driver's phase. The driver bandwidth Δf_x is 0.4 Hz (top row) and 3.2 Hz (bottom row). Note that the maximum amplitude is not always at a phase of 0 or π (i.e. respectively the peaks or the troughs of the slow oscillation). In figure (d), we can also observe that the peak frequency is slightly modulated by the phase of the driver (phase-frequency coupling).

narrow-band driver (in blue) seemed poorly related to the slow oscillation of the raw-signal. To know which of these two slow varying signals best captured the temporal amplitude modulations of the high frequencies, we fitted a model using each of these two drivers. We used the *same* high-pass filtered signal y, i.e. where we removed all frequencies below 16 Hz.

We found that the model with the wide-band driver explained more variance in the high frequencies than the model with the narrow-band driver, that is, that the model better explained the amplitude fluctuations of the high frequencies. As we used the same signal y for both models and as the driver lied in different frequency intervals, it should be noted that the fitting to the slow oscillation visible in Fig. 5(b) and the quality of the fit of the amplitude modulation in the high frequencies were two independent observations. By optimizing for the latter, one observe in Fig. 5b that it leads to a more realistic extraction of the driving neural oscillation.

More generally, our results emphasize the complexity of the choice of parameters in PAC analysis. With our method, we were able to easily compare different parameters, even on non-simulated data, which offered a principled way to set them and helped avoiding misinterpretation due to bad choices. We now investigate the conditional PSDs and the comodulograms estimates on these three datasets.

PSDs and comodulograms estimates of human and rodent neurophysiological recordings

Fig. 6 shows for each signal the PSD (in dB) depending on the driver's phase, estimated through DAR models. The hyper-parameters (p, m) for the rodent striatal, the rodent hippocampal and the human cortical data were $(90, 2)$, $(15, 2)$, and $(24, 2)$, respectively. These parameters were chosen by cross-validation with an exhaustive grid search: $p \in [1, 100]$ and $m \in [0, 3]$. The filtering parameters $(f_x,\ \Delta f_x)$ were chosen to maximize the likelihood as described in previous section, and are respectively $(8.2,\ 3.2)$ Hz, $(5.2,\ 3.2)$ Hz and $(4.0,\ 3.2)$ Hz (Fig. 6 bottom).

For comparison, we also show the PSD obtained with a bandwidth $\Delta f_x = 0.4$ Hz, with the center frequency chosen to maximize the likelihood. We used $(6.4,\ 0.4)$ Hz, $(3.2,\ 0.4)$ Hz and $(2.2,\ 0.4)$ Hz (Fig. 6

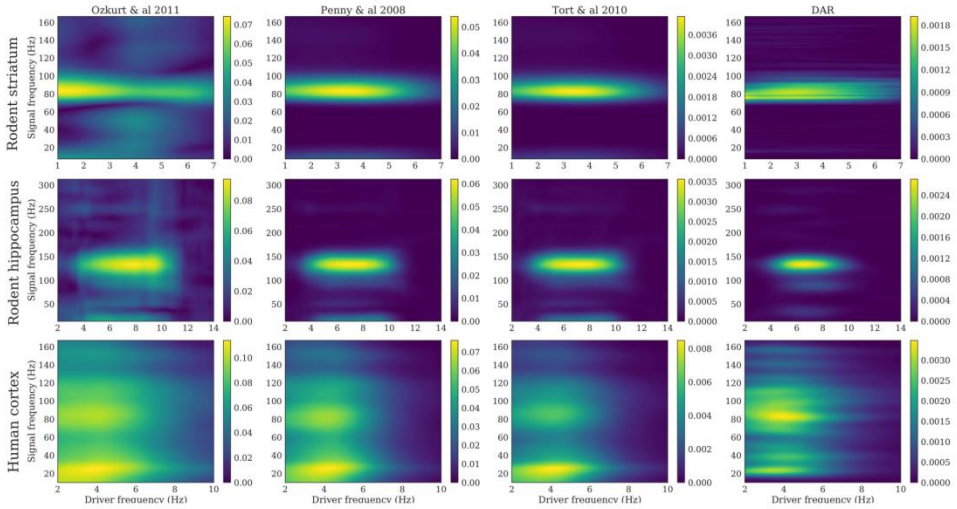


Fig 7. Comodulograms with parameters optimizing the likelihood. Dataset: rodent striatum (top), rodent hippocampus (middle) and human auditory cortex (down). Methods, from left to right: [Ozkurt *et al.* 2011] [Özkurt and Schnitzler, 2011], [Penny *et al.* 2008] [Penny et al., 2008], [Tort *et al.* 2010] [Tort et al., 2010], DAR models (our). The DAR model parameters and the driver bandwidth are chosen to be optimal with respect to the likelihood.

top), respectively.

We observed two kinds of phenomena. In the rodent striatal and hippocampal data, the coupling was mainly concentrated around a given high frequency (i.e. 80 Hz and 125 Hz, respectively), whereas in the human cortical data, the coupling was observed at all frequencies, with a maximum around 20 Hz. The smoothness of the figures depends on the parameter p: a low value leads to a smoother PSD. Please note that the interpretation of the results depends on the filtering parameters.

The two phenomena can also be visualized in Fig. 7, where we plotted comodulograms for all three signals, computed with four different PAC metrics: the mean vector length first proposed in [Canolty et al., 2006] and updated by [Özkurt and Schnitzler, 2011], the parametric model of [Penny et al., 2008] based on GLM, the modulation index of [Tort et al., 2010] that uses the Kullback-Leibler divergence, and our method based on DAR models.

In DAR models, the hyper-parameters used are the cross-validated values described above. With other methods, high frequencies were extracted with a bandwidth Δf_y twice the highest driver frequency used: 14 Hz, 28 Hz, and 20 Hz respectively. For all methods, the drivers were extracted with the optimal bandwidth $\Delta f_x = 3.2$ Hz.

The resulting comodulograms did not look like typically reported comodulograms, as we used a much larger bandwidth (3.2 Hz) than what was found in the literature. With an arbitrary choice of parameters, we could have obtained more classical comodulograms, as shown in Fig. 8. However, these results could be misleading because they suggest a coupling between sinusoidal oscillators, although the corresponding drivers are not perfectly sinusoidal (see Fig. 5(b)).

On these comodulograms, the four methods yielded comparable results as the data we used were very long: 1800, 100, and 730 seconds for the rodent striatal, the rodent hippocampal and the human cortical data, respectively. However, as we will see in the next section, differences emerge between methods when the signals are shorter.

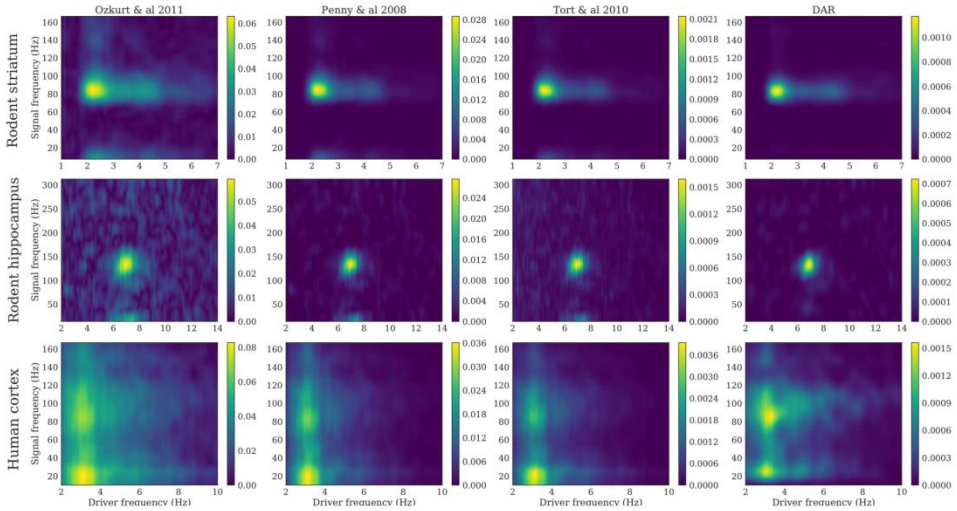


Fig 8. Comodulograms with parameters optimizing the comodulgoram sharpness. Dataset: rodent striatum (top), rodent hippocampus (middle) and human auditory cortex (down). Methods, from left to right: [Ozkurt *et al.* 2011] [Özkurt and Schnitzler, 2011], [Penny *et al.* 2008] [Penny et al., 2008], [Tort *et al.* 2010] [Tort et al., 2010], DAR models (our). The driver bandwidth is chosen to have a well defined maximum in driver frequency: $\Delta f_x = 0.4\,\text{Hz}$. The DAR model parameters are chosen to give similar results than the other methods: $(p, m) = (10, 2)$.

Statistical power and robustness to small sample sizes

Given that DAR models are parametric with a limited number of parameters to estimate, less time samples may be needed to estimate PAC as compared to non-parametric methods. We tested this assumption using simulated signals of varying duration. We computed their comodulograms (as in Fig. 7) and selected the frequencies of maximum coupling. For each duration, we simulated 200 signals, and plotted the 2D histogram showing the fraction of time each frequency pairs corresponded to a maximum. We then compared the same four methods: DAR models with $(p, m) = (10, 1)$, the GLM-based model [Penny et al., 2008], and two non-parametric methods [Tort et al., 2010, Özkurt and Schnitzler, 2011]. Results shown in Fig. 9 show that parametric approaches provided a more robust estimation of PAC frequencies with short signals ($T = 2$ sec) than non-parametric methods.

The robustness to small sample size is a key feature of parametric models, as it significantly improves PAC analysis during shorter experiments. When undertaking a PAC analysis across time using a sliding time window, parametric models should therefore provide more robust PAC estimates. Note that the specific time values in these simulations should not be taken as general guidelines as they depend on the simulation parameters such as the signal-to-noise ratio. However, across all tests, parametric methods consistently provided more accurate results than non-parametric ones.

The amplitude fluctuations of the driver matter

One can note that in DAR models, the driver contains not only the phase of the slow oscillation, but also its amplitude. As the driver is not a perfect sinusoid, its amplitude fluctuates with time. On the contrary, most PAC metrics discard the amplitude fluctuations of the slow oscillation and only consider its phase. To evaluate these two options, we compared two drivers using DAR models: the original driver $x_c(t)$, and the

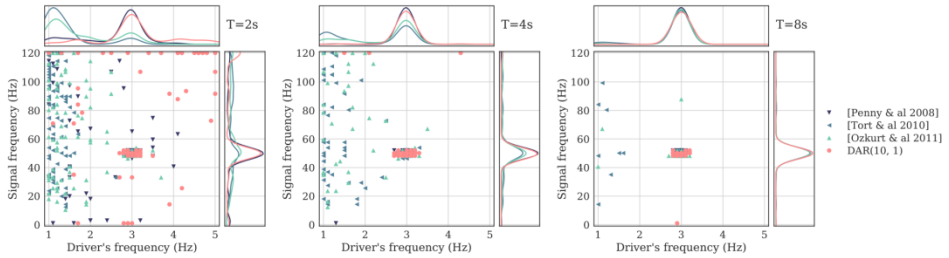


Fig 9. Robustness to small samples. Frequencies of the maximum PAC value, with four methods: a DAR model with $(p, m) = (10, 1)$, the GLM-based model from Penny *et al.* [Penny et al., 2008], and two non-parametric models from Tort *et al.* [Tort et al., 2010] and Ozkurt *et al.* [Özkurt and Schnitzler, 2011]. Each point corresponds to one signal out of 200. Kernel density estimates are represented above and at the right of each scatter plot. From left to right, the simulated signals last $T = 2$, 4 and 8 seconds. The signals are simulated with a PAC between 3 Hz and 50 Hz. The DAR models correctly estimate this pair of frequency even with a short signal length, as well as the GLM-based metric [Penny et al., 2008], while the two other metrics [Tort et al., 2010, Özkurt and Schnitzler, 2011] are strongly affected by the small length of the signals, and do not estimate the correct pair of frequency.

normalized driver $\tilde{x}_c(t) = x_c(t)/|x_c(t)|$. This normalized driver only contains the phase information, as in most traditional PAC metrics. Using cross-validation, we compared the log-likelihood of four fitted models, and found a difference always in favor of the non-normalized driver $x_c(t)$, as it can be visualized in S3 Fig. This result shows that the coupling phenomenon is associated with amplitude fluctuations, a kind of phase/amplitude-amplitude coupling, as it was previously observed in [van Wijk et al., 2015]. Indeed, the GLM parametric method [Penny et al., 2008] was improved when taking into account the amplitude of the slow oscillation. Here, we use our generative model framework to provide an easy comparison tool through the likelihood, to validate this neuroscientific insight from the signals.

Directionality and delay estimation

In this section, we report the results of the directionality estimates using both simulations and neurophysiological signals. It is noteworthy that in DAR models, we arbitrarily call driver the slow oscillation although the model makes no assumption on the directionality of the coupling.

Simulations and comparisons

To validate the directionality estimations approach, we simulated signals with $T = 1024$ time points (4 seconds). We introduced a delay between the slow oscillation and the amplitude modulation of the fast oscillation, and verified that our method could correctly estimate the delays. We did not use a perfect sinusoidal driver, as the delay would only end up in a phase shift, and would not change the strength of the PAC. We used a bandwidth $\Delta f_x = 2$ and a noise level $\sigma_\varepsilon = 0.4$.

We compared our method, using a DAR model with $(p, m) = (10, 1)$, to the cross-frequency directionality (CFD) approach described in [Jiang et al., 2015]. This method makes use of the phase slope index (PSI) [Nolte et al., 2008] for PAC estimation. In [Jiang et al., 2015], the PSI over a frequency band F is defined as:

$$\mathrm{PSI}(F) = \Im\Big(\sum_{f \in F} C_{xy}(f)^* C_{xy}(f + \delta_f)\Big) \tag{17}$$

where C_{xy} is the complex coherence, $\Im$ is the imaginary part, and δ_f is the frequency resolution. As the PSI

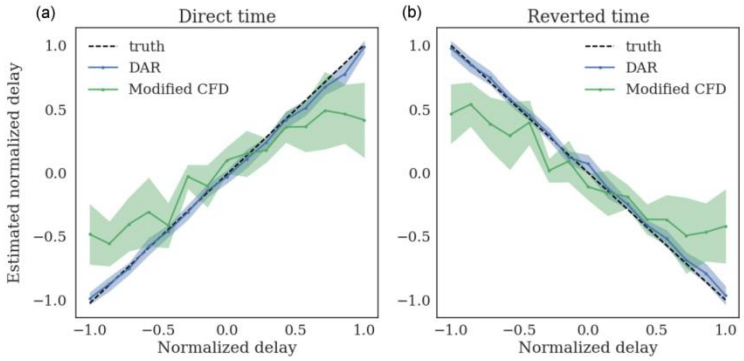


Fig 10. Estimating delays over 20 simulations for each delay. The delays are estimated on both direct time (a) and reverted time (b). The delays are normalized: $\tau = 1$ corresponds to one driver oscillation, i.e. $1/f_x$ sec. The modified CFD shows a bias and only serves for qualitative comparison. The delay estimation based on DAR models correctly estimates the delays, with no visible bias.

was not designed to provide an estimation of the delay, we *modify* it into:

$$\tau_{\mathrm{PSI}}(F) = \frac{\mathrm{PSI}(F)}{2\pi\delta_f n_F} \ , \tag{18}$$

where n_F is the number of frequencies in set F. Note that this delay estimator was correct only if the coherence was almost perfect $C_{xy}(f) \approx 1$ and if the phase slope was small enough to have $\sin(\phi(f+\delta_f) - \phi(f)) \approx \phi(f+\delta_f) - \phi(f)$. As these assumptions are rarely met in practice, we only used this estimator to compare *qualitatively* with our own estimator.

Results presented in Fig. 10 show the mean estimated delay with ± 1 standard deviation computed over 20 simulations. The delay was correctly estimated in both time directions, with no visible bias. As expected, the modified CFD was biased, and showed a higher variance than the DAR-based approach.

Delay estimates of human and rodent data

We applied this delay estimation method on the three experimental recordings, with both direct time and reverted time. Once again, note that the comparison with the CFD [Jiang et al., 2015] is only qualitative, since the CFD was not designed as a delay estimator. The temporal delay was estimated with DAR models with the parameters obtained from the cross-validation selection. The driver we used was obtained with the best filtering parameters found in our grid-search: $(f_x,\ \Delta f_x)$ are respectively $(4.0,\ 3.2)$ Hz, $(8.2,\ 3.2)$ Hz, and $(5.2,\ 3.2)$ Hz.

The results are presented in Fig. 11, where we display the delay estimated with maximum likelihood over multiple DAR models. We also display error bars which correspond to the standard deviation obtained with a block-bootstrap strategy [Carlstein, 1986]. This strategy consists in splitting the signal into $n = 100$ non-overlapping blocks of equal length, to draw at random (with replacement) n blocks, and to evaluate the delay on this new signal. We repeated this process 20 times and computed the standard deviation of the 20 estimated delays. Such strategy is used to estimate empirically the variance of a general statistic from stationary time series.

First, one can observe that the results were qualitatively similar between the two methods. The CFD was biased toward zero for large delays, as was previously shown in the simulations. One can also note that the directionality was not always the same. For the rodent striatal and the human cortical data, the delay was negative and the best model fit happened between the signal y and the *future* driver x. Inversely, in the

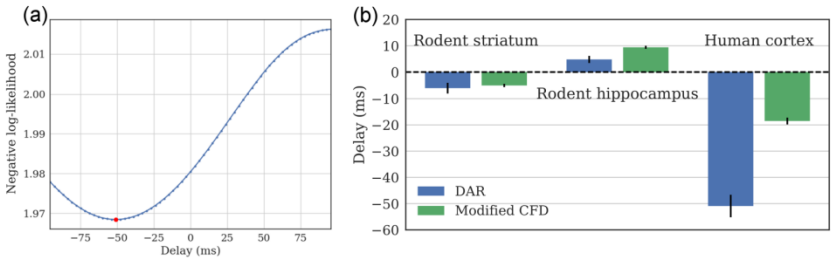


Fig 11. Estimated delays on rodent and human datasets. (a) Negative log likelihood for multiple delays, with the human cortical signal. (b) Optimal delay for the three signals, computed with model selection on DAR models, and with a CFD method modified to provide a delay. The error bars indicate the standard deviation obtained with a block-bootstrap strategy. For the rodent hippocampal data, the delay is positive: the low frequency oscillation precedes the high frequency oscillation. For the human cortical data and the rodent striatal data, the delay is negative: The high frequency oscillation precedes the low frequency oscillation.

rodent hippocampal data, the delay was positive, and the best model fit was between the signal y and the *past* driver x. Further experiments need to be performed to better understand the origin of such delays but we demonstrate here the use of DAR models to estimate them.

Note that our method does not select the delay that leads to the maximum PAC, which would not be statistically valid. On the contrary, we select the delay that leads to the best fit of the model on the data. This approach is more rigorous since we maximize the variance explained by our model, and not the effect of interest.

Discussion

Cross-frequency coupling (CFC) and phase-amplitude coupling (PAC) more specifically have been proposed to play a fundamental role in neural processes ranging from the encoding, maintenance and retrieval of information [Buzsáki, 2010, Jensen and Colgin, 2007, Lisman and Jensen, 2013, Axmacher et al., 2010, Fell and Axmacher, 2011, Kaplan et al., 2014, Hyafil et al., 2015], to large-scale communication across neural ensembles [Canolty and Knight, 2010, Jirsa and Müller, 2013, Khan et al., 2013, Florin and Baillet, 2015]. While a steady increase in observations of PAC in neural data has been seen, how to best detect and quantify such phenomena remains difficult to settle. We argue that a method using DAR models, as described here, is rich enough to capture the time-varying statistics of brain signals in addition to provide efficient inference algorithms. These non-linear statistical models are probabilistic, allowing the estimation of their goodness of fit to the data, and allowing for an easy and fully controlled comparison across models and parameters. In other words, they offer a unique principled data-driven model selection approach, an estimation strategy of phase/amplitude-amplitude coupling based on the approximation of the actual signals, a better temporal resolution of dynamic PAC and the estimation of coupling directionality.

One of the main features of PAC estimation through our method is the ability to compare models or parameters on non-synthetic data. On the contrary, traditional PAC metrics *cannot* be compared on non-synthetic data, and two different choices of parameters can lead to different interpretations. There is no legitimate way to decide which parameter shall be used with empirical data using traditional metrics. The likelihood of the DAR model that can be estimated on left-out data offers a rigorous solution to this problem.

We presented here results on both simulated signals and empirical neurophysiological signals. The simulations gave us an illustration of the phenomenon we want to model, and helped us understand how to visualize a fitted DAR model. They also served a validation purpose for the bandwidth selection approach that we performed on real data. Using the data-driven parameter selection on non-synthetic signals, we

showed how to choose sensible parameters for the filtering of the slow oscillation. All empirical signals are different, and it was for example reported in the neuroscience literature that peak frequencies vary between individuals [Haegens et al., 2014] and that this should not be overlooked in the analysis of the data. The parameter selection based on fitted DAR models makes it possible to fit parameters on individual datasets. Our results also shed light on the asymmetrical and wide-band properties of the slow oscillation, which could denote crucial features involved in cognition [Cole and Voytek, 2017].

The second novelty of our method stands in considering the amplitude fluctuations of the slow oscillation in the PAC measure and not only its phase. Using the rodent and human data, we showed that the instantaneous amplitude of the slow oscillation influences the coupling in PAC, as it was previously suggested in [van Wijk et al., 2015]. The amplitude information should therefore not be discarded as it is done by existing PAC metrics. For instance, the measure of alpha/gamma coupling reported during rest [Osipova et al., 2008, Roux et al., 2013] should incorporate alpha fluctuations when studied in the context of visual tasks [Voytek et al., 2010], as an increase of alpha power is often concomitant with a decrease of gamma power [Fries et al., 2001]. The comparison between DAR models considering or not these low-frequency power fluctuations would inform on the nature of the coupling: purely phase-amplitude, or rather phase/amplitude-amplitude. In Tort *et al.* [Tort et al., 2008], both theta power changes and modulation of theta/gamma PAC were reported in rats having to make a left or right decision to find a reward in a maze. The use of our method could decipher whether the changes in coupling were related to the changes in power, informing on the underlying mechanisms of decision-making. Moreover, as our method models the entire spectrum simultaneously, a phase-frequency coupling could potentially be captured in our models. Therefore, our method is not limited to purely phase-amplitude coupling, and extends the traditional CFC analysis.

Furthermore, in those types of experiments, changes in PAC can be very fast depending on the cognitive state of the subject. Therefore, the need for dynamic PAC estimates is growing [Tort et al., 2008]. We showed with simulations that DAR models are more robust than non-parametric methods when estimating PAC on small time samples. This robustness is critical for time-limited experiments and also when analyzing PAC across time in a fine manner, typically when dynamic processes are at play.

Last but not least, likelihood comparison can also be used to estimate the delay between the coupled components, which would give new insights on highly debated questions on the role of oscillations in neuronal communication [Fries, 2005, Bastos et al., 2015]. For example, a delay close to zero could suggest that the low and high frequency components of the coupling might be generated in the same area, whereas a large delay would suggest they might come from different areas. As an alternative interpretation, the two components may come from the same area, but the coupling mechanism itself might be lagged. In this case, a negative delay would suggest that the low frequency oscillation is driven by the high frequency oscillations, whereas a positive delay would suggest that the low frequency oscillation drives the high frequency amplitude modulation. In any case, this type of analysis will provide valuable information to guide further experimental questions.

A recent concern in PAC analysis is that all PAC metrics may detect a coupling even though the signal is not composed of two cross-frequency coupled oscillators [Kramer et al., 2008, Lozano-Soldevilla et al., 2016, Amiri et al., 2016]. It may happen for instance with sharp slow oscillations, described in humans intracranial recordings [Cole et al., 2016]. Sharp edges are known not to be well described by a Fourier analysis, which decomposes the signal in a linear combination of sinusoids. Indeed, such sharp slow oscillations create artificial high frequency activity at each sharp edge, and these high frequencies are thus artificially coupled with the slow oscillations. This false positive detection is commonly referred to as "spurious" coupling [Jensen et al., 2016]. Even though our method does not use filtering in the high frequencies, it does not solve this issue and is affected in the same way as other traditional PAC metrics. Indeed, our work shed light on the wide-band property of the slow oscillations, but DAR models cannot cope with full-band slow oscillations, which contain strong harmonic components in the high frequencies. However, we consider that such "spurious" PAC can also be a relevant feature of a signal, as stated in [Cole et al., 2016]. In their study, they show that abnormal beta oscillations (13-30 Hz) in the basal ganglia and motor cortex underlie some "spurious" PAC, but are actually a strong feature associated with Parkinson's disease. A robust way to disentangle the different mechanisms that lead to similar PAC results remains to be developed.

The method we presented in this paper uses univariate signals obtained invasively in rodents or humans.

As a lot of neurophysiological research uses non-invasive MEG or EEG recordings containing multiple channels, a multivariate analysis could be of high interest. One way to use data from multiple channels is to estimate a single signal using a spatial filter such as in [Cohen, 2017]. Such a method is therefore complementary to univariate PAC metrics like ours which can be applied to the output of the spatial filter. The method from [Cohen, 2017] builds spatial filters that maximize the difference between, say, high-frequency activity that appears during peaks of a low-frequency oscillation *versus* high-frequency activity that is unrelated to the low-frequency oscillation. Again, from the signal obtained with the spatial filter, it is straightforward to adapt most PAC metrics such as our method.

Neurophysiological signals have all the statistical properties to make them a challenge from a signal processing perspective. They contain non-linearities, non-stationarities, they are noisy and they can be long, hence posing important computational challenges. Our method based on DAR models offer novel and more robust possibilities to analyse neurophysiological signals, paving the way for new insights on how our brain functions via spectral interactions using local or distant coupling mechanisms.

Our method is fully available as an open source python package that comes with documentation, tests, and examples: https://pactools.github.io.

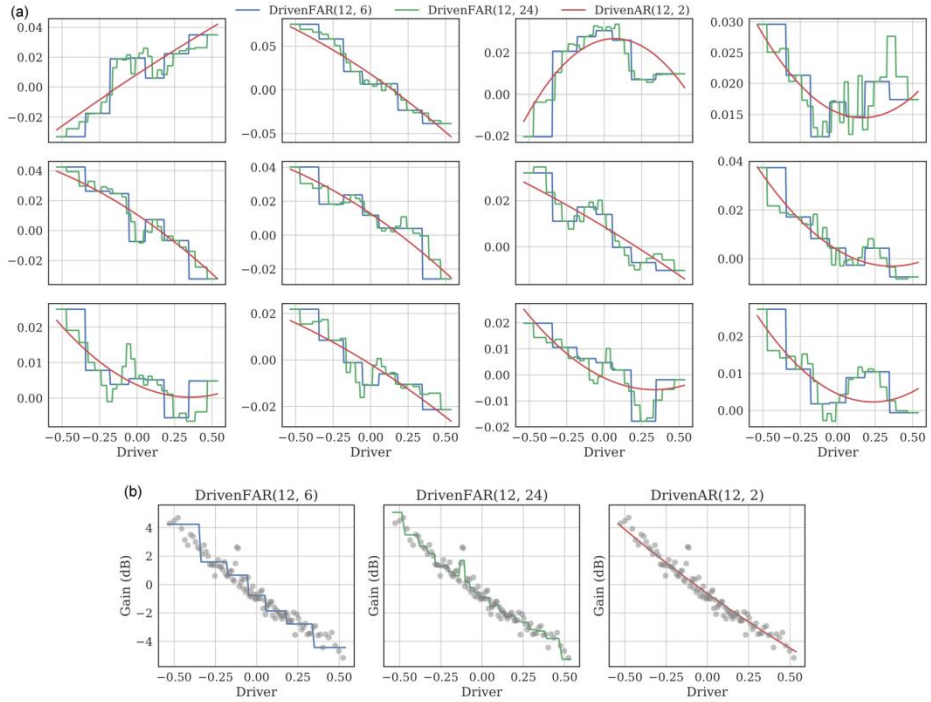
Acknowledgments

We would like to thank Khodagholy *et al.* [Khodagholy et al., 2015] for sharing rodent hippocampal LFP sample data and Canolty *et al.* [Canolty et al., 2006] for sharing human cortical ECoG sample data.

This work was supported by the ERC Starting Grant SLAB ERC-YStG-676943 to Alexandre Gramfort, the ERC Starting Grant MindTime ERC-YStG-263584 to Virginie van Wassenhove, the ANR-16-CE37-0004-04 AutoTime to Valérie Doyère and Virginie van Wassenhove, and the Paris-Saclay IDEX NoTime to Valérie Doyère, Alexandre Gramfort and Virginie van Wassenhove,

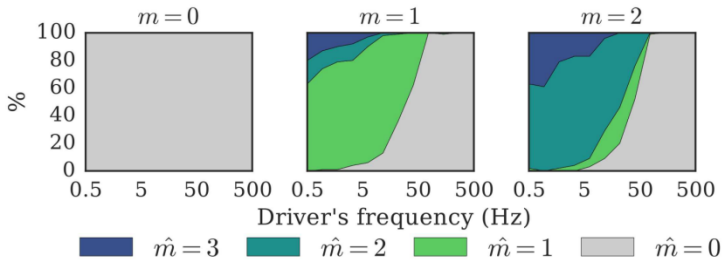
Supporting Information

S1 Fig. Validating the polynomial basis for DAR. (a) AR coefficients $a_i(x)$ as functions of the driver x. (b) Innovation variance $\sigma^2(x)$ as functions of the driver x. The gray circles are the mean squared-residual over 100 bins of the driver values. We compared three parametrizations of these functions: Two staircase functions with respectively $s = 7$ and $s = 25$ steps, and one polynomial function with order $m = 2$, as used in the rest of this work. For a piecewise constant AR parametrization with s-steps, we divided the driver's values into s equally distributed bins, then we fitted s independent linear AR models on the time samples of each bin. This approach is more general than the polynomial approach, yet each AR model uses only T/s samples, whereas the polynomial approach in the DAR model uses all the T samples. The DAR model is therefore more robust. The polynomial parametrization also uses fewer coefficients since a low order m is sufficient. The models are fitted on the human cortical signal, using $p = 12$. We see that an order-2 polynomial is sufficient to approximate the trend of both the AR coefficients and the innovation variance, as estimated by the staircase functions, yet the polynomial uses much fewer parameters.



S1 Fig. Validating the polynomial basis for DAR.

S2 Fig. Model selection with the Bayesian information criterion (BIC). We simulated 100 signals from 100 DAR models with $p = 10$ and $m \in [0, 1, 2]$. We then estimated new DAR models on these signals, with p ranging from 1 to 20, and m ranging from 0 to 3. We selected $\hat{p}$ and $\hat{m}$ that minimized the BIC. The graphs show the proportion of runs that lead to each value of $\hat{m}$. The hyper-parameter m is correctly estimated in most cases if the driver's frequency is not too high ($f_x < 50\,\text{Hz}$). The hyper-parameter p is correctly estimated at ± 2 in most cases (not shown).

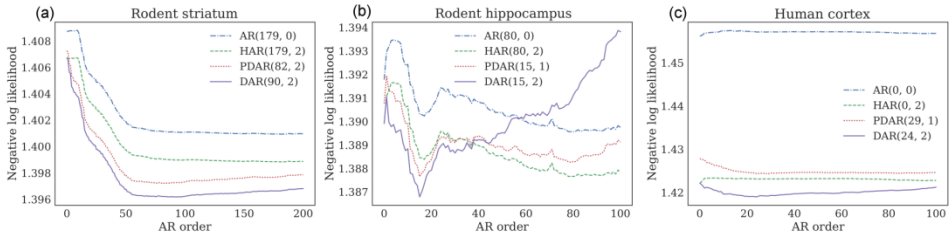


S2 Fig. Model selection with the Bayesian information criterion (BIC).

S3 Fig. Model selection with cross-validation. Cross-validation on our three dataset, (a) rodent striatum, (b) rodent hippocampus, and (c) human auditory cortex, to select the best model and the best parameters. Splitting the signal in half, we fitted the models on the first half, and evaluated the model likelihood on the second half. We compared four different models on a grid of parameter $p \in [0, 100 - 200]$ and $m \in [0, 3]$:

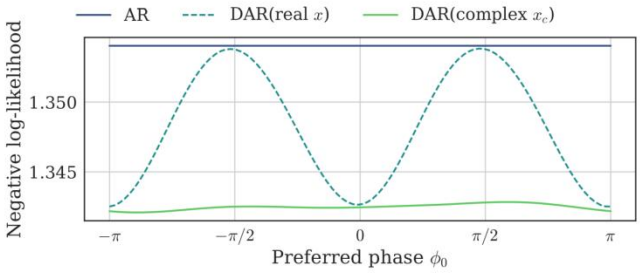
1. AR: a linear AR model
2. Heteroskedastic AR (HAR): an hybrid model between a linear AR model and a DAR model, where the innovation variance σ^2 is driven by x_c, but the AR coefficients are constant in time.
3. Phase DAR (PDAR): a DAR model, with a normalized driver: $x_c/|x_c|$. In this way, we only consider the phase of the slow oscillation, as in most PAC metrics.
4. DAR: a DAR model, where both the innovation variance σ^2 and the AR coefficients are driven by x_c.

The figures present the negative log likelihood (lower is better) by time sample. Each line corresponds to a given model with its best parameter m. The legend shows which orders (p, m) are the best for each model. One can observe that the curves of negative log-likelihood are not convex, yet they exhibit rather clear minima used to define the optimal paramaters.



S3 Fig. Model selection with cross-validation.

S4 Fig. Removing the bias with a complex-valued driver. We simulated a signal as described in the Methods section, introducing a phase difference ϕ_0 in the modulation. For each value of ϕ_0, we fitted three different models, and compared their negative log-likelihood by time sample (the lower the better): an AR, a DAR with a real driver, and a DAR with a complex driver. The parameters were set to $p = 10$ and $m = 1$. A bias is visible around $\phi_0 = \pm\pi/2$, since the real-valued driver DAR model does not fit better than the AR model. As expected, this bias disappears when we update the model to a complex-valued driver DAR model.



S4 Fig. Removing the bias with a complex-valued driver.

References

Akaike, 1998. Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, pages 199–213. Springer.

Amiri et al., 2016. Amiri, M., Lina, J.-M., Pizzo, F., and Gotman, J. (2016). High frequency oscillations and spikes: separating real hfos from false oscillations. *Clinical Neurophysiology*, 127(1):187–196.

Arlot et al., 2010. Arlot, S., Celisse, A., et al. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys*, 4:40–79.

Aru et al., 2015. Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., Singer, W., and Vicente, R. (2015). Untangling cross-frequency coupling in neuroscience. *Current opinion in neurobiology*, 31:51–61.

Axmacher et al., 2010. Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., and Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences*, 107(7):3228–3233.

Axmacher et al., 2006. Axmacher, N., Mormann, F., Fernández, G., Elger, C. E., and Fell, J. (2006). Memory formation by neuronal synchronization. *Brain research reviews*, 52(1):170–182.

Baccalá and Sameshima, 2001. Baccalá, A. L. and Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics*, 84(6):463–474.

Barnett et al., 2009. Barnett, L., Barrett, A. B., and Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for gaussian variables. *Physical review letters*, 103(23):238701.

Bastos et al., 2015. Bastos, A. M., Vezoli, J., and Fries, P. (2015). Communication through coherence with inter-areal delays. *Current opinion in neurobiology*, 31:173–180.

Berman et al., 2012. Berman, J. I., McDaniel, J., Liu, S., Cornew, L., Gaetz, W., Roberts, T. P., and Edgar, J. C. (2012). Variable bandwidth filtering for improved sensitivity of cross-frequency coupling metrics. *Brain connectivity*, 2(3):155–163.

Besserve et al., 2010. Besserve, M., Schölkopf, B., Logothetis, N. K., and Panzeri, S. (2010). Causal relationships between frequency bands of extracellular signals in visual cortex revealed by an information theoretic analysis. *Journal of computational neuroscience*, 29(3):547–566.

Bigdely-Shamlo et al., 2015. Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., and Robbins, K. A. (2015). The prep pipeline: standardized preprocessing for large-scale eeg analysis. *Frontiers in neuroinformatics*, 9:16.

Bollerslev, 1986. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

Bonnefond et al., 2017. Bonnefond, M., Kastner, S., and Jensen, O. (2017). Communication between brain areas based on nested oscillations. *eNeuro*, 4(2):ENEURO–0153.

Bragin et al., 1995. Bragin, A., Jandó, G., Nádasdy, Z., Hetke, J., Wise, K., and Buzsáki, G. (1995). Gamma (40-100 hz) oscillation in the hippocampus of the behaving rat. *The Journal of Neuroscience*, 15(1):47–60.

Bruns, 2004. Bruns, A. (2004). Fourier-, hilbert-and wavelet-based signal analysis: are they really different approaches? *Journal of Neuroscience methods*, 137(2):321–332.

Bruns and Eckhorn, 2004. Bruns, A. and Eckhorn, R. (2004). Task-related coupling from high-to low-frequency signals among visual cortical areas in human subdural recordings. *International Journal of Psychophysiology*, 51(2):97–116.

Buzsáki, 2006. Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford University Press.

Buzsáki, 2010. Buzsáki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. *Neuron*, 68(3):362–385.

Canolty et al., 2006. Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Berger, M. S., Barbaro, N. M., and Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science*, 313(5793):1626–1628.

Canolty and Knight, 2010. Canolty, R. T. and Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in cognitive sciences*, 14(11):506–515.

Carlstein, 1986. Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics*, pages 1171–1179.

Chan and Tong, 1986. Chan, K. S. and Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7(3):179–190.

Chavez et al., 2006. Chavez, M., Besserve, M., Adam, C., and Martinerie, J. (2006). Towards a proper estimation of phase synchronization from time series. *Journal of Neuroscience methods*, 154(1):149–160.

Chehelcheraghi et al., 2017. Chehelcheraghi, M., van Leeuwen, C., Steur, E., and Chie, N. (2017). A neural mass model of cross frequency coupling. *PLoS ONE 12(4): e0173776*.

Cohen, 2017. Cohen, M. X. (2017). Multivariate cross-frequency coupling via generalized eigendecomposition. *eLife*, 6:e21792.

Cole et al., 2016. Cole, S. R., Peterson, E. J., van der Meij, R., de Hemptinne, C., Starr, P. A., and Voytek, B. (2016). Nonsinusoidal oscillations underlie pathological phase-amplitude coupling in the motor cortex in parkinson's disease. *bioRxiv*, page 049304.

Cole and Voytek, 2017. Cole, S. R. and Voytek, B. (2017). Brain oscillations and the importance of waveform shape. *Trends in Cognitive Sciences*.

Colgin et al., 2009. Colgin, L. L., Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., Moser, M.-B., and Moser, E. I. (2009). Frequency of gamma oscillations routes flow of information in the hippocampus. *Nature*, 462(7271):353–357.

Dahlhaus, 1996. Dahlhaus, R. (1996). On the kullback-leibler information divergence of locally stationary processes. *Stochastic Processes and their Applications*, 62(1):139–168.

Dallérac et al., 2017. Dallérac, G., Graupner, M., Knippenberg, J., Martinez, R. C. R., Tavares, T. F., Tallot, L., El Massioui, N., Verschueren, A., Höhn, S., Bertolus, J. B., et al. (2017). Updating temporal expectancy of an aversive event engages striatal plasticity under amygdala control. *Nature Communications*, 8:13920.

Dupré la Tour et al., 2017. Dupré la Tour, T., Grenier, Y., and Gramfort, A. (2017). Parametric estimation of spectrum driven by an exogenous signal. In *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, pages 4301–4305. IEEE.

Dvorak and Fenton, 2014. Dvorak, D. and Fenton, A. A. (2014). Toward a proper estimation of phase–amplitude coupling in neural oscillations. *Journal of Neuroscience methods*, 225:42–56.

Engle, 1982. Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Fell and Axmacher, 2011. Fell, J. and Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nature reviews. Neuroscience*, 12(2):105.

Florin and Baillet, 2015. Florin, E. and Baillet, S. (2015). The brain's resting-state activity is shaped by synchronized cross-frequency coupling of neural oscillations. *NeuroImage*, 111:26–35.

Fries, 2005. Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, 9(10):474–480.

Fries, 2015. Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron*, 88(1):220–235.

Fries et al., 2001. Fries, P., Reynolds, J. H., Rorie, A. E., and Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, 291(5508):1560–1563.

Granger, 1988. Granger, C. W. (1988). Some recent development in a concept of causality. *Journal of econometrics*, 39(1):199–211.

Grenier, 1983. Grenier, Y. (1983). Time-dependent arma modeling of nonstationary signals. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 31(4):899–911.

Grenier, 2013. Grenier, Y. (2013). Estimating an ar model with exogenous driver. Technical Report 2013D007, Telecom ParisTech.

Grenier and Omnes-Chevalier, 1988. Grenier, Y. and Omnes-Chevalier, M.-C. (1988). Autoregressive models with time-dependent log area ratios. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(10):1602–1612.

Haegens et al., 2014. Haegens, S., Cousijn, H., Wallis, G., Harrison, P. J., and Nobre, A. C. (2014). Inter-and intra-individual variability in alpha peak frequency. *Neuroimage*, 92:46–55.

Haggan and Ozaki, 1981. Haggan, V. and Ozaki, T. (1981). Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika*, 68(1):189–196.

Haufe et al., 2010. Haufe, S., Tomioka, R., Nolte, G., Müller, K.-R., and Kawanabe, M. (2010). Modeling sparse connectivity between underlying brain sources for EEG/MEG. *Bio. Eng., IEEE Trans.*, 57(8):1954–1963.

Heusser et al., 2016. Heusser, A. C., Poeppel, D., Ezzyat, Y., and Davachi, L. (2016). Episodic sequence memory is supported by a theta- gamma phase code. *Nature neuroscience*.

Huth et al., 2016. Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., and Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600):453–458.

Hyafil, 2015. Hyafil, A. (2015). Misidentifications of specific forms of cross- frequency coupling: three warnings. *Frontiers in Neuroscience*, 9.

Hyafil et al., 2015. Hyafil, A., Giraud, A.-L., Fontolan, L., and Gutkin, B. (2015). Neural cross-frequency coupling: Connecting architectures, mechanisms, and functions. *Trends in Neurosciences*, 38(11):725–740.

Jensen and Colgin, 2007. Jensen, O. and Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in cognitive sciences*, 11(7):267–269.

Jensen et al., 2016. Jensen, O., Spaak, E., and Park, H. (2016). Discriminating valid from spurious indices of phase- amplitude coupling. *eneuro*, pages ENEURO–0334.

Jiang et al., 2015. Jiang, H., Bahramisharif, A., van Gerven, M. A., and Jensen, O. (2015). Measuring directionality between neuronal oscillations of different frequencies. *Neuroimage*, 118:359–367.

Jirsa and Müller, 2013. Jirsa, V. and Müller, V. (2013). Cross-frequency coupling in real and virtual brain networks. *Frontiers in computational neuroscience*, 7:78.

Kaplan et al., 2014. Kaplan, R., Bush, D., Bonnefond, M., Bandettini, P. A., Barnes, G. R., Doeller, C. F., and Burgess, N. (2014). Medial prefrontal theta phase coupling during spatial memory retrieval. *Hippocampus*, 24(6):656–665.

Kay et al., 2008. Kay, K. N., Naselaris, T., Prenger, R. J., and Gallant, J. L. (2008). Identifying natural images from human brain activity. *Nature*, 452(7185):352–355.

Khan et al., 2013. Khan, S., Gramfort, A., Shetty, N., M., K., Ganesan, S., Moran, J., Lee, S., Gabrieli, J., Tager-Flusberg, H., Joseph, R., Herbert, M., Hämäläinen, M., and Kenet, T. (2013). Local and long-range functional connectivity is reduced in concert in autism spectrum disorders. *Proc. Natl. Acad. Sci.*

Khodagholy et al., 2015. Khodagholy, D., Gelinas, J. N., Thesen, T., Doyle, W., Devinsky, O., Malliaras, G. G., and Buzsáki, G. (2015). Neurogrid: recording action potentials from the surface of the brain. *Nature neuroscience*, 18(2):310–315.

Kikuchi et al., 2017. Kikuchi, Y., Attaheri, A., Wilson, B., Rhone, A. E., Nourski, K. V., Gander, P. E., Kovach, C. K., Kawasaki, H., Griffiths, T. D., Howard III, M. A., et al. (2017). Sequence learning modulates neural responses and oscillatory coupling in human and monkey auditory cortex. *PLoS biology*, 15(4):e2000219.

Kramer et al., 2008. Kramer, M. A., Tort, A. B., and Kopell, N. J. (2008). Sharp edge artifacts and spurious coupling in EEG frequency comodulation measures. *Journal of Neuroscience methods*, 170(2):352–357.

Lachaux et al., 1999. Lachaux, J.-P., Rodriguez, E., Martinerie, J., Varela, F. J., et al. (1999). Measuring phase synchrony in brain signals. *Human brain mapping*, 8(4):194–208.

Lakatos et al., 2005. Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., and Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of neurophysiology*, 94(3):1904–1911.

Lisman and Jensen, 2013. Lisman, J. E. and Jensen, O. (2013). The theta-gamma neural code. *Neuron*, 77(6):1002–1016.

Lozano-Soldevilla et al., 2016. Lozano-Soldevilla, D., ter Huurne, N., and Oostenveld, R. (2016). Neuronal oscillations with non-sinusoidal morphology produce spurious phase-to-amplitude coupling and directionality. *Frontiers in Computational Neuroscience*, 10.

Mahan et al., 2015. Mahan, M. Y., Chorn, C. R., and Georgopoulos, A. P. (2015). White noise test: detecting autocorrelation and nonstationarities in long time series after arima modeling. In *Proceedings 14th Python in Science Conference (Scipy 2015), Austin, TX*.

Malerba and Kopell, 2013. Malerba, P. and Kopell, N. (2013). Phase resetting reduces theta–gamma rhythmic interaction to a one-dimensional map. *Journal of mathematical biology*, pages 1–26.

Naselaris et al., 2011. Naselaris, T., Kay, K. N., Nishimoto, S., and Gallant, J. L. (2011). Encoding and decoding in fmri. *Neuroimage*, 56(2):400–410.

Nolte et al., 2010. Nolte, G., Ziehe, A., Krämer, N., Popescu, F., and Müller, K.-R. (2010). Comparison of granger causality and phase slope index. In *NIPS Causality: Objectives and Assessment*, pages 267–276.

Nolte et al., 2008. Nolte, G., Ziehe, A., Nikulin, V., Schlögl, A., Krämer, N., Brismar, T., and Müller, K.-R. (2008). Robustly estimating the flow direction of information in complex physical systems. *Physical Review Letters*, 100:234101.

Osipova et al., 2008. Osipova, D., Hermes, D., and Jensen, O. (2008). Gamma power is phase-locked to posterior alpha activity. *PloS one*, 3(12):e3990.

Özkurt and Schnitzler, 2011. Özkurt, T. E. and Schnitzler, A. (2011). A critical note on the definition of phase– amplitude cross-frequency coupling. *Journal of Neuroscience methods*, 201(2):438–443.

Park et al., 2015. Park, H., Ince, R. A., Schyns, P. G., Thut, G., and Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Current Biology*, 25(12):1649–1653.

Penny et al., 2008. Penny, W., Duzel, E., Miller, K., and Ojemann, J. (2008). Testing for nested oscillation. *Journal of Neuroscience methods*, 174(1):50–61.

Proakis and Manolakis, 1996. Proakis, J. G. and Manolakis, D. G. (1996). *Digital signal processing 3 rd edition*. Prentice Hall.

Quiroga et al., 2002. Quiroga, R. Q., Kraskov, A., Kreuz, T., and Grassberger, P. (2002). Performance of different synchronization measures in real data: a case study on eeg signals. *Physical Review E*, 65(4):041903.

Roux et al., 2013. Roux, F., Wibral, M., Singer, W., Aru, J., and Uhlhaas, P. J. (2013). The phase of thalamic alpha activity modulates cortical gamma-band activity: evidence from resting- state meg recordings. *Journal of Neuroscience*, 33(45):17827–17835.

Schreiber, 2000. Schreiber, T. (2000). Measuring information transfer. *Physical review letters*, 85(2):461.

Schwarz et al., 1978. Schwarz, G. et al. (1978). Estimating the dimension of a model. *Ann. Stat.*, 6(2):461–464.

Shirvalkar et al., 2010. Shirvalkar, P. R., Rapp, P. R., and Shapiro, M. L. (2010). Bidirectional changes to hippocampal theta–gamma comodulation predict memory for recent spatial episodes. *Proceedings of the National Academy of Sciences*, 107(15):7054–7059.

Spyers-Ashby et al., 1998. Spyers-Ashby, J., Bain, P., and Roberts, S. (1998). A comparison of fast fourier transform (fft) and autoregressive (ar) spectral estimation techniques for the analysis of tremor data. *Journal of neuroscience methods*, 83(1):35–43.

Sweeney-Reed et al., 2014. Sweeney-Reed, C. M., Zaehle, T., Voges, J., Schmitt, F. C., Buentjen, L., Kopitzki, K., Esslinger, C., Hinrichs, H., Heinze, H.-J., Knight, R. T., et al. (2014). Corticothalamic phase synchrony and cross-frequency coupling predict human memory formation. *Elife*, 3:e05352.

Tong and Lim, 1980. Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 245–292.

Tort et al., 2010. Tort, A. B., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *J. Neurophysiol.*, 104(2):1195–1210.

Tort et al., 2008. Tort, A. B., Kramer, M. A., Thorn, C., Gibson, D. J., Kubota, Y., Graybiel, A. M., and Kopell, N. J. (2008). Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a t-maze task. *Proc. Natl. Acad. Sci.*, 105(51):20517–20522.

Tort et al., 2007. Tort, A. B., Rotstein, H. G., Dugladze, T., Gloveli, T., and Kopell, N. J. (2007). On the formation of gamma-coherent cell assemblies by oriens lacunosum-moleculare interneurons in the hippocampus. *Proceedings of the National Academy of Sciences*, 104(33):13490–13495.

Valdés-Sosa et al., 2005. Valdés-Sosa, P. A., Sánchez-Bornot, J. M., Lage-Castellanos, A., Vega-Hernández, M., Bosch-Bayard, J., Melie-García, L., and Canales-Rodríguez, E. (2005). Estimating brain functional connectivity with sparse multivariate autoregression. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 360(1457):969–981.

van Wijk et al., 2015. van Wijk, B., Jha, A., Penny, W., and Litvak, V. (2015). Parametric estimation of cross-frequency coupling. *Journal of neuroscience methods*, 243:94–102.

Voytek et al., 2010. Voytek, B., Canolty, R. T., Shestyuk, A., Crone, N., Parvizi, J., and Knight, R. T. (2010). Shifts in gamma phase–amplitude coupling frequency from theta to alpha over posterior cortex during visual tasks. *Frontiers in human neuroscience*, 4:191.

Wibral et al., 2013. Wibral, M., Pampu, N., Priesemann, V., Siebenhühner, F., Seiwert, H., Lindner, M., Lizier, J. T., and Vicente, R. (2013). Measuring information-transfer delays. *PloS one*, 8(2):e55809.

References

Adhikari, B. M., Goshor, E. S., Lamichhane, B., & Dhamala, M. (2013). Temporal-order judgment of audiovisual events involves network activity between parietal and prefrontal cortices. *Brain Connectivity, 3(5)*, 536-545.

Akam, T., & Kullmann, D. M. (2014). Oscillatory multiplexing of population codes for selective communication in the mammalian brain. *Nat Rev Neurosci., 15*, 112-122.

Albrigth, T. D. (1989). Centrifugal directional bias in the middle temporal visual area (MT) of the macaque. *Visual Neuroscience, 2(2)*, 177-188.

Allport, D. A. (1968). Phenomenal simultaneity and the perceptual moment. *British Journal of Psychology, 59*, 395-406.

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends Cogn Sci., 16(7)*, 390-398.

Arnold, D. H., Clifford, C. W., & Wenderoth, P. (2001). Asynchronous processing in vision: color leads motion. *Current Biology, 11(8)*, 596-600.

Asai, T., & Kanayama, N. (2012). "Cutaneous Rabbit" Hops Toward a Light: Unimodal and Cross-Modal Causality on the Skin. *Front Psychol., 3:427*.

Bachmann, T. (2013). Neurobiological mechanisms behind the spatiotemporal illusions of awareness used for advocating prediction or postdiction. *Front Psychol., 3(593)*, doi: 10.3389/fpsyg.2012.00593.

Baldo, M. V., Kihara, A. H., & Namba, J. (2002). Evidence for an attentional component of the perceptual misalignment between moving and ?ashing stimuli. *Perception, 31(1)*, 17-30.

Baldo, M. V., Klein, S. A., Khurana, B., & Nijhawan, R. (1995). Extrapolation or attention shift? *Nature, 378(6557)*, 565-566.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., . . . Halgren, E. (2006). Top-down facilitation of visual recognition. *Proc Natl Acad Sci USA, 103(2)*, 449-454.

Barascud, N., Pearce, M. T., Griffiths, T. D., Friston, K. J., & Chait, M. (2016). Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proceedings of the National Academy of Sciences of the United States of America, 113(5)*.

Bastos, A. M., Vezoli, J., & Fries, P. (2015). Communication through coherence with inter-areal delays. *Current Opinion in Neurobiology, 31*, 173-180.

Battelli, L., Pascual-Leone, A., & Cavanagh, P. (2007). The "when" pathway of the right parietal lobe. *Trends Cogn Sci., 11(5)*, 204-210.

Baumgarten, T. J., Schnitzler, A., & Lange, J. (2015). Beta oscillations define discrete perceptual cycles in the somatosensory domain. *Proc, 112(39)*, 12187-192.

Baumgarten, T. J., Schnitzler, A., & Lange, J. (2016). Prestimulus alpha power influences tactile temporal perceptual discrimination and confidence in decisions. *Cereb Cortex, 26(3)*, 891-903.

Bear, A., & Bloom, P. (2016). A simple task uncovers a postdictive illusion of choice. *Psychol Sci., 27(6)*, 914-922.

Bechlivanidis, C., & Lagnado, D. A. (2016). Time reordered: Causal perception guides the interpretation of temporal order. *Cognition, 146*, 58-66.

Beckers, G., & Zeki, S. (1995). The consequences of inactivating areas V1 and V5 on visual motion perception. *Brain, 1*, 49-60.

Beer, A. L., Plank, T., & Greenlee, M. W. (2011). Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex. *Experimental Brain Research, 213(299)*.

Benwell, C. S., Learmonth, G., Miniussi, C., Harvey, M., & Thut, G. (2015). Non-linear effects of transcranial direct current stimulation as a function of individual baseline performance: evidence from biparietal tDCS influence on lateralized attention bias. *Cortex, 69*, 152-165.

Bernasconi, F., Grivel, J., Murray, M. M., & Spierer, L. (2010). Interhemispheric coupling between the posterior sylvian regions impacts successful auditory temporal order judgment. *Neuropsychologia, 48(9)*, 2579-85.

Bernasconi, F., Manuel, A. L., Murray, M. M., & Spierer, L. (2011). Pre-stimulus beta oscillations within left posterior sylvian regions impact auditory temporal order judgment accuracy. *International Journal of Psychophysiology, 79*, 244-248.

Binder, M. (2015). Neural correlates of audiovisual temporal processing--comparison of temporal order and simultaneity judgments. *Neuroscience, 300*, 432-447.

Blankenburg, F., Ruff, C. C., Deichmann, R., Rees, G., & Driver, J. (2006). The cutaneous rabbit illusion affects human primary sensory cortex somatotopically. *PLoS Biology, 4(3)*, e69 doi: 10.1371/journal.pbio.0040069.

Boenke, L. T., Deliano, M., & Ohl, F. W. (2009). Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. *Exp Brain Res., 198(2-3)*, 233-44.

Bonnefond, M., & Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Current Biology, 22(20)*, 1969-74.

Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience, 28*, 157-189.

Bragin, A., Jando, G., Nadasdy, Z., Hetke, J., Wise, K., & Buzsaki, G. (1995). Gamma (40-100 Hz) oscillation in the hippocampus of the behaving rat. *Journal of Neuroscience, 15(1)*, 47-60.

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics, 62(8)*, 1572-1595.

Brown, J. W., & Braver, T. S. (2005). Learned Predictions of Error Likelihood in the Anterior Cingulate Cortex. *Science, 307(5712)*, 1118-121.

Buehner, M. J., & May, J. (2003). Rethinking temporal contiguity and the judgement of causality: Effects of prior knowledge, experience, and reinforcement procedure. *The Quaterly Journal of Experimental Psychology, 56A(5)*, 865-890.

Bueti, D., Lasaponara, S., Cercignani, M., & Macaluso, E. (2012). Learning about Time: Plastic Changes and Interindividual Brain Differences. *Neuron, 75(4)*, 725-737.

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci., 29(24)*, 7869-76.

Bushara, K. O., Grafman, J., & Hallett, M. (2001). Neural correlates of auditory-visual stimulus onset asynchrony detection. *J of Neurosci., 21(1)*, 300-304.

Buzsaki, G. (2006). *Rhythms of the Brain.* Oxford Scholarship.

Buzsaki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science, 304*, 1926-1929.

Canavier, C. C. (2015). Phase-resetting as a tool of information transmission. *Current Opinion in Neurobiology, 31*, 206-213.

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., . . . Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science, 313(5793)*, 1626-1628.

Cecere, R., Rees, G., & Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Current Biology, 25(2)*, 231-235.

Chakravarthi, R., & VanRullen, R. (2012). Conscious updating is a rythmic process. *PNAS, 109(26)*, 10599-604.

Chaumon, M., & Busch, N. (2014). Prestimulus Neural Oscillations Inhibit Visual Perception via Modulation of Response Gain. *J Cogn Neurosci., 26(11)*, 2514-29.

Chechlacz, M., Gillebert, C. R., Vangkilde, S. A., Petersen, A., & Humphreys, G. W. (2015). Structural variability within frontoparietal networks and individual differences in attentional functions: an approach using the theory of visual attention. *J Neurosc., 35(30)*, 10647-658.

Choi, H., & Scholl, B. J. (2006). Perceiving causality after the fact: postdiction in the temporal dynamics of causal perception. *Perception, 35(3)*, 385-399.

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics, 1362*, 48-55.

Coull, J. T., & Nobre, A. C. (2008). Dissociating explicit timing from temporal expectation with fMRI. *Curr Opin Neurobiol, 18(2)*, 137-144.

Coull, J. T., Cheng, R.-K., & Meck, W. H. (2011). Neuroanatomical and Neurochemical Substrates of Timing. *Neuropsychopharmacology, 36*, 3-25.

Cravo, A. M., Santos, K. M., Reyes, M. B., Caetano, M. S., & Claessens, P. M. (2015). Visual causality judgments correlate with the phase of alpha oscillations. *J Cogn Neurosci.*, 1-8.

Crowe, D. A., Zarco, W., Bartolo, R., & Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *J Neurosci., 34(36)*, 11972-83.

Davis, B., Christie, J., & Rorden, C. (2009). Temporal order judgments activate temporal parietal junction. *J of Neurosci., 29(10)*, 3182-88.

Dehaene, S. (1993). Temporal oscillations in human perception. *Psychological Scie, 4(4)*, 264-270.

Deisz, R. A., & Prince, D. A. (1989). Frequency-dependent depression of inhibition in guinea-pig neocortex in vitro by GABAB receptor feed-back on GABA release. *Journal of Physiology, 412*, 513-541.

Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer: the where and when of consciousness in the brain. *Behavioral and Brain Sciences, 15(2)*, 183-247.

Dubois, J. (2016). Building a Science of Individual Differences from fMRI. *Trends in Cognitive Science, 20(6)*, 425-443.

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(33)*.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science, 287*, 2036.

Efron, R. (1970). The minimum duration of a perception. *Neuropsychologia, 8(1)*, 57-63.

Efron, R. (1970). The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia, 8(1)*, 37-55.

Eimer, M., Forster, B., & Vibell, J. (2005). Cutaneous saltation within and across arms: a new measure of the saltation illusion in somatosensation. *Percept Psychophys, 67(3)*, 458-68.

Engel, A. K., & Fries, P. (2010). Beta-band oscillations--signalling the status quo? *Curr Opin Neurobiol, 20(2)*, 156-165.

Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn Sci., 5(1)*, 16-25.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic prediction: oscillations and synchrony in top-down processing. *Nat Rev Neurosci., 2(10)*, 704-16.

Enns, J. T., & Lollo, V. D. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4(9)*, 345-352.

Ergenoglu, T., Demiralp, T., Bayraktaroglu, Z., Ergen, M., Beydagi, H., & Uresin, Y. (2004). Alpha rhythm of the EEG modulates visual detection performance in humans. *Cognitive Brain Research, 20(3)*, 376-383.

Flach R., H. P. (2006). The cutaneous rabbit revisited. *J Exp Psycholo Hum Percept Perform., 32(3)*, 717-732.

Fontolan, L., Morillon, B., Liegeois-Chauvel, C., & Giraud, A.-L. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nature Communications, 5(4694)*.

Fraisse, P. (1984). Perception and estimation of time. *Ann. Rev. Psychol, 35*, 1-36.

Franciotti, R., Brancucci, A., Penna, D., Onofrj, M., & Tommasi, L. (2011). Neuromagnetic responses reveal the cortical timing of audiovisual synchrony. *Neuroscience*.

Freeman, E. D., Ipser, A., Palmbaha, A., Paunoiu, D., Brown, P., Lambert, C., . . . Driver, J. (2013). Sight and sound out of synch: Fragmentation and renormalisation of audiovisual integration and subjective timing. *Cortex, 49(10)*, 2875-2887.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci., 9(10)*, 474-80.

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science, 291(5508)*, 1560-563.

Friston, K. J. (2005). A theory of cortical responses. *Phil. Trans. R. Soc. Biol Sci., 360(1456)*, 815-836.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat Neurosci., 7(7)*, 773-778.

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clin Neurophysiol., 120(3)*, 453-463.

Gauthier, B., & Wassenhove, V. v. (2016). Time is not space: core computations and domain-specific networks for mental travels. *Journal of Neuroscience, 36(47)*, 11891-903.

Geldard, F. A. (1977). Cutaneous stimuli, vibratory and saltatory. *Journal of Investigative Dermatology, 69(1)*, 83-87.

Geldard, F. A. (1982). Saltation in somesthesis. *Psychological Bulletin, 92(1)*, 136-175.

Geldard, F. A., & Sherrick, C. E. (1972). The cutaneous "rabbit": a perceptual illusion. *Science., 178(4057)*, 178-9.

Genovesio, A., Tsujimoto, S., & Wise, S. P. (2009). Feature- and Order-Based Timing Representations in the Frontal Cortex. *Neuron, 63(2)*, 254-266.

Gibbon, J., & Rutschmann, R. (1969). Temporal order judgment and reaction time. *Science, 165(3891)*, 413-415.

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience, 15*, 511-517.

Goldreich, D., & Tong, J. (2013). Prediction, postdiction, and perceptual length contraction: a Bayesian low-speed prior captures the cutaneous rabbit and related illusions. *Front Psychol., 4*, 221.

Gramfort, A., Luissi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., . . . Hamalainen, M. S. (2014). MNE software for processing MEG and EEG data. *Neuroimage, 86*, 446-460.

Hadjikhani, N., Liu, A. K., Dale, A. M., Cavanagh, P., & Tootell, R. B. (1998). Retinotopy and color sensitivity in human visual cortical area V8. *Nature Neuroscience, 1(3)*, 235-241.

Haegens, S., Nácher, V., Luna, R., Romo, R., & Jensen, O. (2011). α-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *PNAS, 108(48)*, 19377-382.

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bäuml, K.-H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *NeuroImage, 37(4)*, 1465-73.

Hari, R. (1995). Illusory directional hearing in humans. *Neuroscience Letters, 189*, 29-30.

Harris, L., Harrar, V., Jaekl, P., & Kopinska, A. (2010). Space and time in perception and action. In B. K. Romy Nijhawan (Ed.). Cambridge university Press.

Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010). Phonological decisions require both the left and right supramarginal gyri. *PNAS, 107(8)*, 16494-499.

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc Natl Acad Sci USA, 109(49)*, 20095-20100.

Heron, J., Whitaker, D., McGraw, P. V., & Horoshenkov, K. V. (2007). Adaptation minimizes distance-related audiovisual delays. *Journal of Vision, 7(5)*, 1-8.

Herrmann, C. S., Munk, M. H., & Engel, A. K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Science, 8(8)*, 347-355.

Heusser, A. C., Poeppel, D., Ezzyat, Y., & Davachi, L. (2016). Episodic sequence memory is supported by a theta-gamma phase code. *Nat Neurosci.*, 1-7.

Hilbert, M. (2012). Toward a synthesis of cognitive biases: how noisy information processing can bias human decision making. *Psychological Bulletin, 138(2)*, 211-237.

Hirsh, I. J., & Jr., C. E. (1961). Perceived order in different sense modalities. *J Exp Psychol, 62(5)*, 423-432.

Hsieh, L.-T., Ekstrom, A. D., & Ranganath, C. (2011). Neural oscillations associated with item and temporal order maintenance in working memory. *Journal of Neuroscience, 31(30)*, 10803-10.

Hubbard, T. L. (2014). The flash-lag effect and related mislocalizations: findings, properties, and theories. *Psychological Bulletin, 140(1)*, 303-38.

Iemi, L., Chaumon, M., Crouzet, S. M., & Busch, N. A. (2017). Spontaneous neural oscillations bias perception by modulating baseline excitability. *Journal of Neuroscience, 37(4)*, 807-819.

Ikumi, N. (2016). *The effect of cognitive factors on cross-modal synchrony perception.* Ph.D. dissertation, Universitat Pompeu Fabra.

Ikumi, N., & Soto-Faraco, S. (2014). Selective attention modulates the direction of audio-visual temporal recalibration. *PLoS BiolS One, 9(7)*.

Ipser, A. (2014). *Individual differences in audiovisual integration and timing.* Ph.D. dissertation, Department of Psychology, City University London.

Ipser, A., Agolli, V., Bajraktari, A., Al-Alawi, F., Djaafara, N., & Freeman, E. D. (2017). Sight and sound persistently out of synch: stable individual differences in audiovisual synchronisation revealed by implicit measures of lip-voice integration. *Scientific Reports, 7(46413)*.

Ishigami, Y., & Phillips, D. P. (2008). Effect of stimulus hemifield on free-field auditory saltation. *Hearing Research, 241(1-2)*, 97-102.

Jensen, O., & Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *TiCS, 11(7)*, 267-269.

Jensen, O., & Mazaheri, A. (2010). Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. *Front Hum Neurosci., 4(186)*.

Jensen, O., Til, B. G., Bergmann, O., & Bonnefond, M. (2014). Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends Neurosci., xx*, xx.

Jones, B., & Huang, Y. L. (1982, Jan). Space-time dependencies in psychophysical judgment of extent and duration: Algebraic models of the tau and kappa effects. *Psychological Bulletin, 91(1)*, 128-142.

Kaganovich, N. S. (2016). Electrophysiological correlates of individual differences in perception of audiovisual temporal asynchrony. *Neuropsychologia, 86*, 119-130.

Kambe, J., Kakimoto, Y., & Araki, O. (2015). Phase reset affects auditory-visual simultaneity judgment. *Cogn Neurodyn*.

Kamitani, Y., & Shimojo, S. (1999). Manifestation of scotomas created by transcranial magnetic stimulation of human visual cortex. *Nature Neuroscience, 2*, 767-771.

Kanabus, M., Szelag, E., Rojek, E., & Pöppel, E. (2002). Temporal order judgement for auditory and visual stimuli. *Acta Neurobiol. Exp., 62*, 263-270.

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nat Rev Neurosci., 12*, 232-242.

Kanai, R., Bahrami, B., & Rees, G. (2010). Human parietal cortex structure predicts individual. *Current Biology, 20(18)*, 1626-1630.

Keane, B., Spence, M., Yarrow, K., & Arnold, D. (2015). Perceptual confidence demonstrates trial-by-trial insight into the precision of audio-visual timing encoding. *Consciousness and Cognition, 38*, 107-117.

Keetels, J. V. (2010). Perception of intersensory synchrony: a Tutorial review. *Attention, Perception, \& Psychophysics, 72(4)*, 871-884.

Kelly, S. P., Gomez-Ramirez, M., & Foxe, J. J. (2009). The strength of anticipatory spatial biasing predicts targetdiscrimination at attended locations: a high-densityEEG study. *European Journal of Neuroscience, 30(11)*, 2224-2234.

Khoei, M. A., Masson, G. S., & Perrinet, L. U. (2017). The Flash-Lag Effect as a Motion-Based Predictive Shift. *PLOS Computational Biology*, 1-31.

Khuu, S. K., Kidd, J. C., & Badcock, D. R. (2011). The influence of spatial orientation on the perceived path of visual saltatory motion. *J Vis., 11(9)*, 1-16.

Khuu, S. K., Kidd, J. C., & Errington, J. A. (2010). The effect of motion adaptation on the position of elements in the visual saltation illusion. *Journa of Vision, 10(12):19*, 1-14.

Kidd, J. C., & Hogben, J. H. (2004). Quantifying the auditory saltation illusion: an objective psychophysical methodology. *J Acoust Soc Am., 116(2)*, 1116-1125.

Kikuchi, Y., Attaheri, A., Wilson, B., Rhone, A. E., Nourski, K. V., Gander, P. E., . . . Petkov, C. I. (2017). Sequence learning modulates neural responses and oscillatory coupling in human and monkey auditory cortex. *PLOS Biology*.

Kilgard, M. P., & Merzenich, M. M. (1995). Anticipated stimuli across skin. *Nature, 373*, 663.

King, A. J. (2005). Multisensory integration: strategies for synchronization. *Current Biology, 15(9)*, R339-R341.

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn Sci., 16*, 606-17.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res Rev., 53(1)*, 63-88.

Kopinska, A., & Harris, L. R. (2004). Simultaneity constancy. *Perception, 33*, 1049-60.

Kösem, A., Gramfort, A., & Wassenhove, V. v. (2014). Encoding of event timing in the phase of neural oscillations. *NeuroImage, 92*, 274-84.

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron, 93*.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends Cogn Sci., 12(2)*, 72-79.

Lackner, J. R., & Teuber, H.-L. (1973). Alterations. *Neuropsychologia, 11*, 409-415.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science, 320(5872)*, 110-113.

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J Neurophysiol., 94(3)*, 1904-911.

Landau, A. N., Schreyer, H. M., Pelt, S. v., & Fries, P. (2015). Distributed attention is implemented through theta-rhythmic gamma modulation. *Curr Biol., 25*, 2332-37.

Lange, J., Halacz, J., Dijk, H. v., Kahlbrock, N., & Schnitzler, A. (2012). Fluctuations of Prestimulus Oscillatory Power Predict Subjective Perception of Tactile Simultaneity. *Cereb Cortex, 22(11)*, 2564-74.

Lange, J., Oostenveld, R., & Fries, P. (2013). Reduced occipital alpha power indexes enhanced excitability rather than improved visual perception. *J Neurosci., 33(7)*, 3212-20.

Lewandowska, M., Bekisz, M., Szymaszek, A., Wrobel, A., & Szelag, E. (2008). Towards electrophysiological correlates of auditory perception of temporal order. *Neuroscience Letters, 437(2)*, 139-143.

Lewis, R., & Noppeney, U. (2010). Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. *J Neurosci., 30(37)*, 12329-12339.

Limbach, K., & Corballis, P. M. (2016). Prestimulus alpha power influences response criterion in a detection task. *Psychophysiology, 53(8)*, 1154-64.

Linares, D., & Holcombe, A. O. (2014). Differences in perceptual latency estimated from judgments of temporal order, simultaneity and duration are inconsistent. *i-Perception, 5(6)*, 559-571.

Lisman, J. E., & Idiart, M. A. (1995). Storage of 7 +/- 2 short-term memories in oscillatory subcycles. *Science, 267(5203)*, 1512-1515.

Lisman, J. E., & Jensen, O. (2013). The theta-gamma neural code. *Neuron, 77*, 1002-1016.

Liu, J., & Newsome, W. T. (2005). Correlation between Speed Perception and Neural Activity in the Middle Temporal Visual Area. *Journal of Neuroscience, 25(3)*, 711-722.

Lockhead, G. R., Johnson, R. C., & Gold, F. M. (1980). Saltation through the blind spot. *Perception & Psychophysics, 27(6)*, 545-549.

Lutz, A., & Thompson, E. (2003). Neurophenomenology integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies, 10(9-10)*, 31-52.

Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2002). Dynamic brain source of visual evoked responses. *Science, 295(5559)*, 690-694.

Maniscalco, B., Peters, M. A., & Lau, H. (2016). Heuristic use of perceptual evidence leads to dissociation between performance and metacognitive sensitivity. *Attention, Perception, & Psychophysics, 78(3)*, 923-37.

Marshall, T. R., Bergmann, T. O., & Jensen, O. (2015). Frontoparietal structural connectivity mediates the top-down control of neuronal synchronization associated with selective attention. *PLoS Biol, 13(10)*, 1-17.

Marshuetz, C. (2005). Order information in working memory: an integrative review of evidence from brain and behavior. *Psychological Bulletin, 131(3)*, 323-339.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *J Neurosci., 29(9)*, 2725-32.

Matthews, W. J., & Meck, W. H. (2014). Time perception: the bad news and the good. *WIREs Cog Sci., 5(4)*, 429-446.

Maus, G. W., Ward, J., Nijhawan, R., & Whitney, D. (2012). The perceived position of moving objects: transcranial magnetic stimulation of area MT1 reduces the flash-lag effect. *Cerebral Cortex, 23*, 241-47.

McDonald, J. J., Teder-Sälejärvi, W. A., F. D., & Hillyard, S. A. (2005). Neural basis of auditory-induced shifts in visual time-order perception. *Nat Neurosci., 8(9)*, 1197-1202.

Mehta, M. R., Lee, A. K., & Wilson, M. A. (2002). Role of experience and oscillations in transforming a rate code into a temporal code. *Nature, 417*, 741-746.

Michalareas, G., Vezoli, J., Pelt, S. v., Schoffelen, J.-M., Kennedy, H., & Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron, 89(2)*, 384-397.

Milton, A., & Pleydell-Pearce, C. W. (2016). The phase of pre-stimulus alpha oscillations influences the visual perception of stimulus timing. *NeuroImage, 133*, 53-61.

Miyazaki, M., Hirashima, M., & Nozaki, D. (2010). The "Cutaneous Rabbit" Hopping out of the Body. *J Neurosci., 30(5)*, 1856-60.

Miyazaki, M., Hiroshi Kadota, K. S., Takeuchi, S., Sekiguchi, H., Aoyama, T., & Kochiyama, T. (2016). Dissociating the neural correlates of tactile temporal order and simultaneity judgements. *Scientific Reports, 6(23323)*.

Morrone, M. C., Tosetti, M., Fiorentini, D. M., Cioni, G., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nat, 3(12)*, 1322-28.

Mueller, S., Wang, D., Fox, M. D., Yeo, B. T., Sepulcre, J., Sabuncu, M. R., . . . Liu, H. (2013). Individual variability in functional connectivity architecture of the Human brain. *Neuron, 77(3)*, 586-595.

Myers, N. E., Walther, L., Wallis, G., Stokes, M. G., & Nobre, A. C. (2015). Temporal dynamics of attention during encoding versus maintenance of working memory: complementary views from event-related potentials and alpha-band oscillations. *Journal of Cognitive Neuroscience, 27(3)*, 492-508.

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 USA, 106(23)*, 9169-73.

Naveh-Benjamin, M. (1990). Coding of temporal order information: an antomatic process? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(1)*, 117-126.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370(6487)*, 256-257.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends Cogn Sci., 6(9)*, 387-393.

Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H.-J., & Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *J Neurosci., 27(42)*, 11431-11441.

O'Keefe, J., & Recce, M. L. (1995). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus, 3(3)*, 317-330.

Orban, G., Berkes, P., Fiser, J., & Lengyel, M. (2016). Neural variability and sampling-based probabilistic representations in the visual cortex. *Neuron, 92*, 530-543.

Palva, J. M., Palva, S., & Kaila, K. (2005). Phase synchrony among neuronal oscillations in the human cortex. *J Neuro., 25(15)*, 3962-3972.

Palva, S., & Palva, J. M. (2011). Functional Roles of Alpha-Band Phase Synchronization in Local and Large-Scale Cortical Networks. *Frontiers in Psychology, 2(204)*.

Patel, M., & Chait, M. (2011). Retroactive adjustment of perceived time. *Cognition, 119(1)*, 125-130.

Perrone, J. A., & Thiele, A. (2001). Speed skills: measuring the visual speed analyzing properties of primate MT neurons. *Nature Neuroscience, 4*, 526-532.

Phillips, D. P., & Hall, S. E. (2001). Spatial and temporal factors in auditory saltation. *J Acoust Soc Am., 110(3)*, 1539-47.

Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition, 121*, 147-153.

Pöppel, E. (1970). Excitability Cycles in Central Intermittency. *Psychol. Forsch., 34*, 1-9.

Pöppel, E. (1972). The Study of Time. In H. F. Fraser J.T. (Ed.). Springer.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends Cogn Sci., 1(2)*, 56-31.

Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Phil. Trans. R. Soc. B, 364*, 1887-96.

Roberts, B. M., Hsieh, L.-T., & Ranganath, C. (2013). Oscillatory activity during maintenance of spatial and temporal information in working memory. *Neuropsychologia, 51(2)*, 349-57.

Rohenkohl, G., & Nobre, A. C. (2011). Alpha oscillations related to anticipatory attention follow temporal expectations. *J Neurosci., 31*, 14076-84.

Romei, V., Brodbeck, V., Michel, C., Amedi, A., Pascual-Leone, A., & Thut, G. (2008). Spontaneous fluctuations in posterior alpha-band EEG activity reflect variability in excitability of human visual areas. *Cerebeal Cortex, 18(9)*, 2010-2018.

Romei, V., Gross, J., & Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *Journal of Neuroscience, 30(25)*, 8692-8697.

Roseboom, W., & H.Arnold, D. (2011). Twice Upon a Time Multiple Concurrent Temporal Recalibrations of Audiovisual Speech. *Psychological Science, 22(7)*, 872-877.

Rubinsten, O., Dana, S., Lavro, D., & Berger, A. (2013). Processing ordinality and quantity: ERP evidence of separate mechanisms. *Brain and Cognition, 82*, 201-212.

Sadaghiani, S., & Kleinschmidt, A. (2016). Brain Networks and alpha Oscillations: Structural and Functional Foundations of Cognitive Control. *Trends in Cognitive Science, 20(11)*, 805-817.

Samaha, J., & Postle, B. R. (2015). The speed of alpha-band oscillations predicts the temporal resolution of visual perception. *Curr Biol., 25*, 1-6.

Samaha, J., Bauer, P., Cimaroli, S., & Postle, B. R. (2015). Top-down control of the phase of alpha-band oscillations as a mechanism for temporal prediction. *PNAS, 112(27)*, 8439-44.

Samaha, J., Iemi, L., & Postle, B. R. (2017). Prestimulus alpha-band power biases visual discrimination confidence, but not accuracy. *Conscious Cogn., S1053-8100(16)*, 30430-35.

Sanford, E. C. (1888). Personal Equation. *A J Psychol, 2(1)*, 3-38.

Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., . . . Birbaumer, N. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. *Eur J Neurosci., 22*, 2917-26.

Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cogn Psychol., 47*, 333-366.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci., 32(1)*, 9-18.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Curr Opin Neurobiol, 20(2)*, 172-176.

Scott A. Love, K. P. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PLoS One, 8(1)*, e54798.

Sedley, W., Gander, P. E., Kumar, S., Kovach, C. K., Oya, H., Kawasaki, H., . . . Griffiths, T. D. (2016). Neural signatures of perceptual inference. *eLife, 5(e11476)*.

Sergent, C., Ruff, C. C., Barbot, A., Driver, J., & Rees, G. (2011). Top-down modulation of human early visual cortex after stimulus offset supports successful postcued report. *Journal of Cognitive Neuroscience, 23(8)*, 1921-34.

Sergent, C., Wyart, V., Babo-Rebelo, M., Cohen, L., Naccache, L., & Tallon-Baudry, C. (2013). Cueing attention after the stimulus is gone can retrospectively trigger conscious perception. *Curr Biol., 23*, 150-155.

Servos, P., Osu, R., Santi, A., & Kawato, M. (2002). The Neural Substrates of Biological Motion Perception: an fMRI Study. *Cerebral Cortex, 12(7)*, 772-782.

Sherman, M. T., Kanai, R., Seth, A. K., & VanRullen, R. (2016). Rhythmic influence of top-down perceptual priors in the phase of prestimulus occipital alpha oscillations. *J Cogn Neurosci.*

Sheth, B. R., Nijhawan, R., & Shimojo, S. (2000). Changing objects lead brief flashed ones. *Nature Neuroscience, 3(5)*, 489-495.

Shimojo, S. (2014). Postdiction: its implications on visual awareness, hindsight, and sense of agency. *Front Psychol., 5(196)*, 1-19.

Shimojo, S., & Kamitani, Y. (2001). Sound-induced visual "rabbit". *Journal of Vision, 1(3)*, 478.

Shore, D. I., Hall, S. E., & Klein, R. M. (1998). Auditory saltation: A new measure for an old illusion. *J Acoust Soc Am., 103(6)*, 3730-3733.

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior-entry. *American Psychological Society, 12(3)*, 205-212.

Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? *Neuron, 24(1)*, 49-65.

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: the Colavita effect revisited. *Perception & Psychophysics, 69(5)*, 637-686.

Soga, R., Akaishi, R., & Sakai, K. (2009). Predictive and postdictive mechanisms jointly contribute to visual awareness. *Consciousness and Cognition, 3*, 578-592.

Sokoliuk, R., & VanRullen, R. (2013). The flickering wheel illusion: when alpha rhythms make a static wheel flicker. *J Neurosci., 33(33)*, 13498-504.

Spence, C., & Parise, C. (2010). Prior-entry: a review. *Conscious Cogn., 19(1)*, 364-379.

Spence, C., & Squire, S. (2003). Multisensory integration: maintaining the perception of synchrony. *Curr Biol., 13(13)*, R519-R521.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *J Exp Psychol Gen., 130(4)*, 799-832.

Stefanics, G., Hangya, B., hernadi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *J Neurosci., 30(41)*, 13578-585.

Stein, A. v., & Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *Int J Psychophysiol., 38(3)*, 301-313.

Steinbüchel, N. v., Wittmann, M., Strasburger, H., & Szelag, E. (1999). Auditory temporal-order judgement is impaired in patients with cortical lesions in posterior regions of the left hemisphere. *Neurosci Letters, 264*, 168-171.

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: fundamental issues and a general model. *Attention & Performance*, 630-681.

Stetson, C., Cui, X., Montague, P. R., & Eagleman, D. M. (2006). Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron, 51(5)*, 651-659.

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., . . . Porter, N. R. (2001). When is now? Perception of simultaneity. *Proc R Soc Lond., 268*, 31-38.

Stroud, J. M. (1967). The fine structure of psychological time. *Annals of the New York Academy of Sciences, 138(2)*, 623-631.

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature, 421(6926)*(911).

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Science, 13(9)*, 403-409.

Takahashi T, K. K. (2013). Neural correlates of tactile temporal-order judgment in humans: an fMRI study. *Cereb Cortex, 23(8)*, 1952-64.

Takahashi, T., & Kitazawa, S. (2017). Modulation of illusory reversal in tactile temporal order by the phase of posterior alpha rhythm. *Journal of Neuroscience*, 2899-2915.

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends Cogn Sci., 14(9)*, 400-410.

Thibault, L., Berg, R. v., Cavanagh, P., & Sergent, C. (2016). Retrospective attention gates discrete conscious access to past sensory stimuli. *PLoS One, 11(2)*.

Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci., 26'37)*, 9494-9502.

Timothy J. Buschman, E. K. (2007). Top-down versus bottom-Up control of attention in the prefrontal and posterior parietal cortices. *Science, 315(5820)*, 1860-62.

Titchener, E. B. (1908). Lectures on the elementary psychology of feeling and attention. *New York: Macmillan*.

Tort, A. B., Kramer, M. A., Thorn, C., Gibson, D. J., Kubota, Y., Graybiel, A. M., & Kopell, N. J. (2008). Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a T-maze task. *Proc Natl Acad Sci USA, 105(51)*, 20517-522.

Tort, A. B., Robert W. Komorowski, Manns, J. R., Kopell, N. J., & Eichenbaum, H. (2009). Theta-gamma coupling increases during the learning of item-context associations. *PNAS, 106(49)*, 20942-947.

Trojan, J., Getzmann, S., Möller, J., Kleinböhl, D., & Hölzl, R. (2009). Tactile-auditory saltation: Spatiotemporal integration across sensory modalities. *Neuroscience Letters, 460*, 156-160.

Tsal, Y. (1983). Movement of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 9(4)*, 523-530.

Tünnermann, J., Petersen, A., & Scharlau, I. (2015). Does attention speed up processing? Decreases and increases of processing rates in visual prior entry. *J of Vision, 15(3)*, 1-27.

Tyler, C. W. (1985). Analysis of visual modulation sensitivity. II. Peripheral retina and the role of photoreceptor dimensions. *J Opt Soc Am A., 2(3)*, 393-398.

Ulrich, R. (1987). Threshold models of temporal-order judgments evaluated by a ternary response task. *Perception & Psychophysics, 42(3)*, 224-239.

Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid recalibration to asynchronous audiovisual stimuli. *Journal of Neuroscience, 33*, 14633-637.

van Dijk, H., Schoffelen, J.-M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J Neurosc, 28(8)*, 1816-23.

Van Eijk, R. L., Kohlrausch, A., Juola, J. F., & Par, S. v. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception & Psychophysics, 70(6)*, 955-968.

van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., Togt, C. v., & Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *PNAS, 111(40)*, 14333-14341.

van Wassenhove, V. (2009). Minding time in an amodal representational space. *Phil. Trans. R. Soc. B, 364*, 1815-1830.

van Wassenhove, V. (2016). Temporal cognition and neural oscillations. *Current Opinion in Behavioral Sciences, 8*, 124-130.

van Wassenhove, V. (2017). Defining moments for conscious time and content. *PsyCh Journal*, 1-2.

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience, 10(11)*, 792-802.

VanRullen, R. (2016). Perceptual cycles. *Trends in Cognitive Science, 20(10)*, 723-735.

VanRullen, R., & Dubois, J. (2011). The psychophysics of brain rhythms. *Front Psychol., 2(203)*, 1-10.

VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Science, 7(5)*, 207-213.

Varela, F. J. (2000). Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science. In F. J.-M. Jean Petitot (Ed.). Stanford University Press, Stanford.

Varela, F. J., Toro, A., John, E. R., & Schwartz, E. L. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia, 19(5)*, 675-686.

Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience, 2*, 229-239.

Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Temporal recalibration during asynchronous audiovisual speech perception. *Experimental Brain Research, 181(1)*, 173-181.

Vibell, J., Klinge, C., Zampini, M., Spence, C., & Nobre, A. C. (2007). Temporal order is coded temporally in the brain: early event-related potential latency shifts underlying prior entry in a cross-modal temporal order judgment task. *J Cogn Neurosci., 19(1)*, 109-120.

Vroomen, J., Keetels, M., Gelder, B. d., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cogn Brain Res., 22*, 32-35.

Weiss, K., Hilkenmeier, F., & Scharlau, I. (2013). Attention and the speed of information processing: posterior entry for unattended stimuli instead of prior entry for attended stimuli. *PLoS One, 8(1)*.

Westlye, L. T., Grydeland, H., Walhovd, K. B., & Fjell, A. M. (2011). Associations between regional cortical thickness and attentional networks as measured by the attention network test. *Cerebral Cortex, 21(2)*, 345-356.

White, C. T. (1963). Temporal numerosity and the psychological unit of duration. *Psychological Monographs: General and Applied, 77(12)*, 1-37.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience, 1(8)*, 656-657.

Wiener, M., Matell, M. S., & Coslett, H. B. (2011). Mutliple mechanisms for temporal processing. *Front Integr Neurosci., 5(31)*, 1-3.

Woo, S.-H., Kim, K.-H., & Lee, K.-M. (2009). The role of the right posterior parietal cortex in temporal order judgment. *Brain and Cognition, 69(2)*, 337-343.

Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *Journal of Neuroscience, 20(6)*, RC63.

Wutz, A., & Melcher, D. (2014). The temporal window of individuation limits visual capacity. *Front Psychol., 5(952)*, 1-13.

Wutz, A., Muschter, E., Koningsbruggen, M. G., Weisz, N., & Melcher, D. (2016). Temporal integration windows in neural processing and perception aligned to saccadic eye movements. *Curr Biol., 26*, 1-10.

Wutz, A., Weisz, N., Braun, C., & Melcher, D. (2014). Temporal windows in visual processing: "prestimulus brain state" and "poststimulus phase reset" segregate visual transients on different temporal scales. *J Neurosci., 34(4)*, 1554-65.

Xia, Y., Morimoto, Y., & Noguchi, Y. (2016). Retrospective triggering of conscious perception by an interstimulus interaction. *Journal of Vision, 16(7)*, 3.

Yuan, X., Li, H., Liu, P., Yuan, H., & Huang, X. (2016). Pre-stimulus beta and gamma oscillatory power predicts perceived audiovisual simultaneity. *Int J Psychophysiol., 107*, 29-36.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127(1)*, 3-21.

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Exp Brain Res., 152(2)*, 198-210.

Zatorre, R., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb Cortex, 11(10)*, 946-53.

Zeki, K. M. (1997). A direct demonstration of perceptual asynchrony in vision. *Proc R Soc B, 264*, 393-399.