Temporal Dynamics in the Multisensory Brain

Institute of Neuroscience

Newcastle University

Mark Laing

06 March 2019

Abstract

In this work, I investigate the mechanisms with which the brain manages temporal coherence between sensory signals. An overview of relevant literature is given, and current theories about how sensory signals are combined in brain and behaviour are introduced. Key unknowns about the temporal dynamics of auditory-visual integration are identified and addressed within four investigations. In the first study, I assess whether cues to the onset of a auditory-visual pair affect sensitivity to their temporal asynchrony. It is shown that regularly timed cues shorten the temporal window of integration compared with irregular cues. This demonstrates that attention can affect how sensory signals are bound. In the second experiment, speech-like asynchronous stimuli are presented for an extended duration whilst perceptual simultaneity is monitored. In this manner, the time-course of temporal adaptation is tracked over time. Adaptation occurs when the presented asynchrony is visual-leading, but not when it is auditory-leading. This may suggest that temporal recalibration in the auditory-leading direction is not a consequence of adaptation. In the third investigation, the neural correlates of the time-course of temporal adaptation are measured. Increased activity in frontal and parietal areas occurred during perceptual asynchrony, this replicates previous work and further promotes that these regions provide top-down modulation of the mechanisms of temporal simultaneity. Increased activity is present in the posterior cingulate cortex whilst the brain is maintaining an adapted state, compared with during adaptation. This region may act as a conflict monitor and compensator for temporal asynchrony. Lastly, I investigate the extent to which a highly prevalent inhibitory neurotransmitter affects performance in a multisensory behavioural task. There is a possible correlation between the concentration of gamma-aminobutyric acid in the parietal lobe and the overall strength of integration effects. Finally, the impact and future directions of this work are discussed in the context of current literature.

Dedication

For Kaylee, my brothers, Mam, Dad, and all my friends without whom I could not have completed this work.

Declaration

This thesis is submitted to Newcastle University for the degree of Doctor of Philosophy. All reported research was performed between September 2014 and March 2018, and supervised by Dr Quoc Vuong and Professor Adrian Rees. The contents of this work are wholly my own and have not been previously submitted by me for a degree or any other qualification.

Acknowledgements

Many thanks to Quoc Vuong and Adrian Rees, whose guidance, patience, and understanding have been instrumental to the completion of this work, and to my personal and academic development. Special thanks go to Kamie Lakhani, Afnaan Nadeem, and Hei Iong Tung for their assistance in collection of behavioural data. I would like to thank Jehill Parikh for his assistance with magnetic resonance spectroscopy. I am grateful to Louise Ward, Tim Hodgson, and Dorothy Wallace for performing radiography. I would finally like to thank the Wellcome Trust and the Institute of Neuroscience for awarding me with the studentship that has made this all possible.

List of Tables

List of Figures

List of Abbreviations

ACC anterior cingluate cortex. 88

ANOVA analysis of variance. 44, 48, 60, 61, 65, 68, 69, 95, 98

BOLD blood-oxygen level dependent. 33, 79, 80, 87, 88, 91, 100

DAT dynamic attending theory. 8, 27, 28, 39, 40, 52, 103

DTI diffusion tensor imaging. 32

EEG electroencephalography. 30, 32, 53, 103, 105

fMRI functional magnetic resonance imaging. 32, 34, 38, 73, 74, 77, 81, 105

FWE family-wise error. 80, 83, 84, 85

GABA gamma-aminobutyric acid. 7, 10, 36, 37, 38, 90, 91, 92, 97, 99, 100, 101, 102, 106, 107

GLM general linear model. 79

IFG inferior frontal gyrus. 84, 85, 87, 88, 97

IPS intraparietal sulcus. 83, 84, 85, 87, 88, 92, 97

ISI inter-stimulus interval. 27, 32, 42, 43, 46, 47, 53

MEG magnetoencephalography. 32, 36

MRI magnetic resonance imaging. 97

MRS magnetic resonance spectroscopy. 9, 36, 90, 96, 97

PCC posterior cingulate cortex. 81, 83, 81, 85, 88, 89, 105

PDF probability distribution function. 51

PFC prefrontal cortex. 33, 74, 75, 80, 106

PPC posterior parietal cortex. 33, 83

PSS point of subjective simultaneity. 20, 21, 23, 24, 25, 26, 37, 54, 55, 71, 104

QUEST quantile estimation. 50, 51

SC superior colliculus. 33, 91, 107

SJ simultaneity judgement. 20, 21, 40, 104

SOA stimulus onset asynchrony. 8, 20, 21, 22, 23, 24, 25, 32, 42, 43, 44, 46, 47, 48, 51, 55, 57, 58, 60, 61, 64, 65, 68, 69, 71, 76, 80

STS superior temporal sulcus. 33, 34, 38, 74, 75, 80, 92, 97, 102

TBW temporal binding window. 8, 20, 21, 22, 23, 24, 37, 39, 40, 41, 42, 43, 44, 48, 49, 51, 52, 103, 104

TOJ temporal order judgement. 20, 21, 24, 39, 42, 43, 48, 49, 52, 103, 104

Contents

Chapter 1

Introduction

Our sensory organs are able to sense a fraction of the electromagnetic waves, pressure fluctuations, and chemical interactions that occur in the natural world. An unsolved question in neuroscience is how all of these signals are bound into a single unified percept [1]. At least part of this problem is addressed by multisensory integration, which combines signals between senses to facilitate the detection, discrimination, and binding of sensory information [2]. This thesis addresses unanswered questions about how sensory signals are combined, and focuses on mechanisms that maintain temporal coherence between the senses. This chapter introduces the fundamentals of multisensory integration, the stimulus features that facilitate its mechanisms, and the current understanding of its underlying neural basis.

1.1 Background

A great deal is known about each sense in isolation, but much less is understood about how signals are combined to create perceptual objects with which we can reason and react [3]. When an event emits signals that can be received by more than one sense, we unconsciously analyse the relationship between these signals in order to make inferences about them. This analysis allows us to make more sense of our environment, and the integrated product reveals more about the external event more quickly and accurately than the sum of each sense in isolation [2]. For multisensory stimuli, reaction times can be much faster [4, 5], and detection thresholds can be much lower [6, 7]. For example, in [6] participants were seated before an array of 8 visual displays and corresponding speakers located at 8°, 24°, 40°, and 56°to the

left and right of a central fixation point. The visual displays comprised of four red Light Emitting Diodes (LEDs) and a central green target LED. On target trials, one of the eight visual displays illuminated the green target LED and four red masking LEDs. On non-target trials, one of the eight visual displays illuminated only the four red LEDs. The duration of the illumination was gradually reduced from 100ms to 60ms between blocks in order to maintain a target detection rate between 60% and 70%. An auditory stimulus (white noise burst) was played on target trials from one of the eight speakers. The white noise burst was the same duration as the visual target, but preceded the visual stimulus by 500ms in 50% of trials. Detection of the target stimulus was *only* improved compared to visual only performance when the accompanying auditory stimulus was presented at the same time and location as the target. Here, the presence of an auditory stimulus demonstrably improved performance in a visual task. This is an example of *multisensory enhancement*.

Multisensory integration is especially effective when sensory signals are degraded [8]. The intelligibility of degraded auditory speech is greatly enhanced when visual speech is introduced [9–13]. For example, in [9] participants were read sequences of 25 bisyllabic words from a restricted vocabulary set (6–128 words) and were asked to indicate which words were pronounced by writing them down. Half of the participants were seated facing away from the speaker (auditory-only speech), and half were seated facing the speaker (auditory-visual speech). White noise was presented to the participants during the reading, and the white noise level was changed between readings to achieve the desired signal to noise ratio (which was measured manually with a decibel meter). As expected, participants performed the most poorly (i.e. most words incorrect) during high signal to noise ratio and the best during low signal to noise ratio. Performance was increased for auditory-visual speech vs. auditory-only speech, and this performance gain was especially pronounced at when the signal to noise ratio was high. This relative performance gain for low and high signal to noise ratios is known as *inverse effectiveness*.

Conversely, when multisensory stimuli are temporally, spatially, or semantically misaligned, the multisensory response is suppressed, as the likelihood that the sensory signals originated from the same events is low [14–17]. For example, spatial incongruence would be present if visual and auditory signals appeared to originate from an observers left and right respectively. These signals would not be considered to have originated in the same event, and the multisensory response to these signals is inhibited [8, 18]. An example of a semantic incongruence would be to see a cat but hear *"woof!"*. Again, multisensory responses to these signals would be inhibited because the sensory

components of the stimuli are semantically opposed to our learned associations about the world [19–21]. For example, in [19] participants were asked to identify the contents of 42 briefly-presented, and then rapidly-masked, pictures of natural stimuli [22]. A sound was presented alongside the picture which was white noise, or was either semantically congruent or incongruent with respect to the target picture. Semantically congruent sounds improved participants identification performance, whereas semantically incongruent sounds impaired performance, as compared to performance in the white-noise control condition This is an example of *multisensory depression*.

The relative enhancements in perception afforded by multisensory integration are contingent on mechanisms that can interpret the relative timing, spatial origin, and semantics of cross-modal sensory signals. If an event emits signals that share these amodal properties, the brain is more likely to treat the signals as originating from the same source [23–26]. This is known as the *unity assumption*. I will consider each of these properties in turn in the sections that follow.

1.2 Temporal Factors

In the temporal domain, we expect that signals from the same event will be emitted simultaneously. Therefore, it could be expected that the brain will only apply integrative enhancements to signals that arrive at the same time. However, the brain must be tolerant of delay in accepting which signals to integrate, since the arrival time of each sensory component of a multisensory stimulus is dependent on the distance between the source, the observer, and its speed of propagation. For example, light travels much faster than sound ($300,000,000ms^{-1}$ vs. $330ms^{-1}$) therefore auditory and visual signals emitted by an event at a distance of 20m will be detected by an observer with a delay of 61ms. Additionally, the time taken for signal transduction is shorter for hearing ($\tilde{1}0ms$) than vision ($\tilde{5}0ms$)[27]. As a result of these physical and neural differences, it has been proposed that sensory signals only arrive synchronously to the auditory and visual cortices when an event occurs approximately $10-15m$ from and observer - this distance has been described as the *horizon of simultaneity* [27]. Thus, if the brain is overly strict when evaluating temporal synchrony, the brain might lead to inhibition of multisensory signals for distant events.

In addition to evaluating the relative onset timing of multisensory signals, the brain must evaluate the coherence of these signals over time. Here it is

important to make a distinction between temporal asynchrony and temporal incongruency. Temporal asynchrony refers to a constant delay between sensory streams, such as if the video and audio on a television is out of synchrony, whereas temporal incongruency refers to uncorrelated sensory streams, such as with an overdubbed speaker.

1.2.1 Temporal Synchrony

Behavioural responses to multisensory stimuli are dependent on the correct resolution of temporal synchrony. Response times to multisensory stimuli are affected by delays between sensory signals, with the fastest response times when the stimuli are synchronous [4, 5]. Temporal synchrony also leads to better performance in motion discrimination [28, 29] and target detection [7, 30]. Therefore, some mechanism must evaluate the temporal synchrony between signals of different sensory modalities.

We can interrogate this mechanism by presenting multisensory stimuli with a delay between components in different sensory modalities known as a stimulus onset asynchrony (SOA). We then ask participants to report whether those components occurred at the same time [24, 31]. This is known as a simultaneity judgement (SJ) task. Figure 1.1 displays how the distribution of participants' *"simultaneous"* responses might look for such an experiment. The temporal binding window (TBW) is a quantification of the time-period over which we are likely to perceive multisensory stimuli to be synchronous. The point of subjective simultaneity (PSS) is the mean of this distribution, and is the SOA at which the subject is most likely to perceive the stimuli as synchronous.

An analogous paradigm is known as the temporal order judgement (TOJ) task, in which participants are presented with equivalent stimuli as in the SJ task, but are asked to report whether the flash preceded the tone or not (*see* Figure 1.1). Using these measurements, we can quantify the TBW by taking the range of SOAs over which participants were "uncertain" of the temporal order of the stimuli. This is usually the difference between 25% and 75% thresholds [24, 31].

If the TBW is measured using the SJ task, the results are correlated with, but not equivalent to, results from the TOJ task [31]. On average, the TBW is shorter when measured using the TOJ task [31, 32]. It is possible that the SJ task is simply more subjective than the TOJ task, since the decision criterion for simultaneity is variable between observers. This has led to the opinion

that the TOJ task is better representative of low-level temporal mechanisms than the SJ task [24].

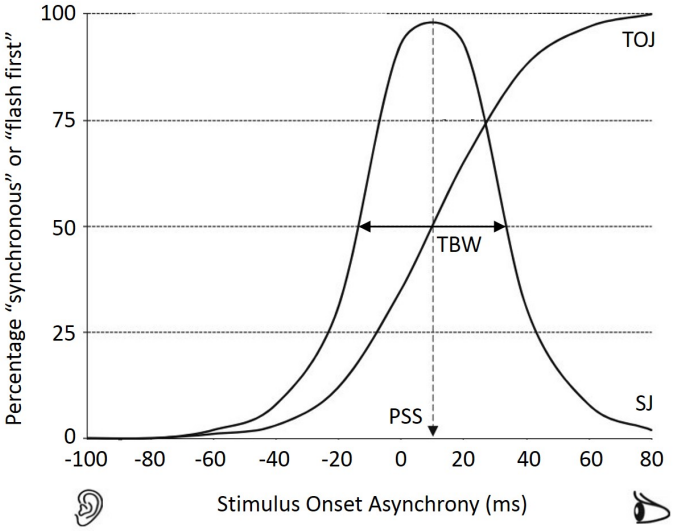

Figure 1.1: The physical delay with which auditory and visual stimuli are presented (SOA), against the likelihood that they are perceived to be synchronous (SJ) and the likelihood that the visual stimulus was perceived before the auditory stimulus (TOJ). At large auditory-leading asynchronies (negative SOAs), participants are unlikely to perceive the stimuli to be synchronous, and unlikely to perceive the visual stimulus before the auditory stimulus. When the delay between sensory components is small, participants are likely to perceive the stimuli to be synchronous, and are equally likely to perceive the visual stimulus or auditory stimulus to occur first. For SJ tasks, participants responses are typically fitted with a Gaussian function, the TBW is usually reported as the duration over which value of the Gaussian function is above 50%, and the PSS is reported as the the mean of the Gaussian function. For TOJ tasks, participants responses are fitted with a psychometric function, the TBW is reported as the duration over which the value of the psychometric function is between 25% and 75%, and the PSS is the SOA at which the value of the psychometric function is 50%. Image adapted from [24].

A consistent property of responses to auditory-visual asynchrony is that the PSS is slightly visual-leading. This means that simultaneity is most likely to be perceived if the visual component is slightly before the auditory component [31–35]. The asymmetry is slight, and usually around 20ms [32]. Since light travels faster than sound, a 20ms visual-leading asynchrony occurs naturally when an event occurs around 6m from an observer. It is most likely that this asymmetry is a learned association about the natural world [24].

A critical finding is that responses to asynchrony are malleable. Within participants, the TBW and PSS may change as a result of stimulus type [31, 32], task demands [31], training [36, 37], and recent exposure to temporal

asynchrony [38, 39]. This implies that the mechanisms which determine the synchrony of multisensory signals operate in a way which is dependent on context.

Stevenson & Wallace [31] carried out a systematic investigation of perceptual synchrony. Participants were presented with stimuli of three different types: transient flashes and beeps, videos of tools (such as a hammer hitting a table), and videos of speech (single syllables). The stimuli were presented with a range of SOAs, between the visual and auditory components. Each stimulus type was then presented 20 times per SOA. Participants indicated whether they perceived the stimuli to be simultaneous at each SOA, and the averaged results could be fitted with a response curve (*see* Figure 1.2). Finally, the TBW for each participant was calculated as the range of SOAs over which their response curve was above 50%. Stevenson et. al (2013) showed that the temporal binding window is longest for speech, shorter for tools, and shortest for transient stimuli [31]. The authors postulated that this effect is related to stimulus complexity, and that the TBW is optimised to handle extended processing times in the primary sensory cortices [31, 33].

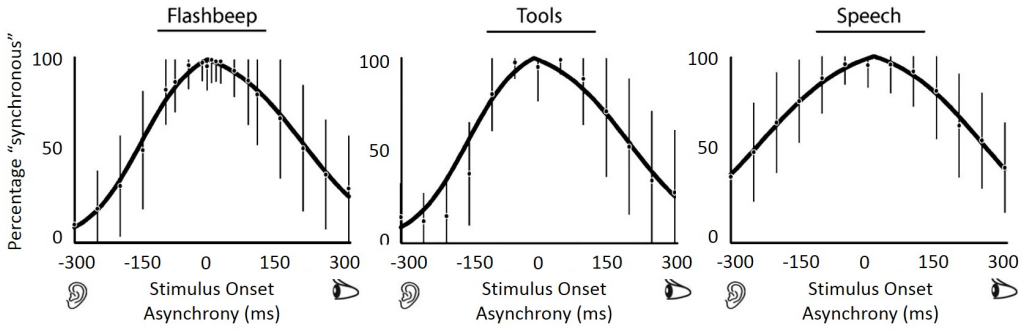

Figure 1.2: Perceptual synchrony, reported as a function of stimulus onset asynchrony (SOA) and stimulus type. Error bars represent the standard deviation between participants. I can see that for transient, flash-beep stimuli (left), the width of the response curve is most narrow. On the right, I can see that the width of the response curve is widest for speech stimuli. The width of these curves directly reflects the width of the TBW as well as the strictness of the mechanism which determines the synchrony of sensory components of a multisensory stimulus. Image adapted from [31]

Powers et al. [37] demonstrated that with feedback training, participants were able to narrow their TBW, thereby reducing their tolerance to sensory asynchrony. In their study, participants completed five 1-hour training sessions on sequential days before their final assessment. In each training session, participants were presented with beep-flash stimuli with various SOAs, and asked to judge whether the beep and flash occurred simultaneously. After submitting their response, participants were presented with the accuracy of their judgement: a yellow smiley face and the word *Correct!* when correct,

or a red sad face with the word *Incorrect* when incorrect. In the test sessions, stimuli and task demands were exactly as in the training sessions, with the exception that a) no feedback was provided, and b) a larger range of SOAs were presented to participants. Participants completed test sessions before and after training. Training significantly affected the width of the TBW, which was narrowed by 40% on average.

Fujisaki et al. [38], demonstrated that the mechanisms responsible for perceptual synchrony can *recalibrate* to favour asynchronies to which the observer has been previously exposed. Before each testing session, the authors presented participants with beeps and flashes with various SOAs (auditory first 235ms, 0ms, visual first 235ms) for three minutes. This is known as the "adaptation phase". Before each trial, participants viewed a "re-adapting" stimulus with the same SOA as in adaptation phase for 10s. Test stimuli were then presented at a range of SOAs. Participants reported whether the test stimuli occurred at the same time or not. After the adaptation phase, participants were more likely to perceive the test stimuli to be synchronous when the modality order of the SOA was the same direction as the modality order of the adaptation stimulus SOA (i.e. visual first or auditory first). This led to a shift in the point of subjective simultaneity (PSS) for the test stimuli in the direction of the asynchrony of the adaptation stimuli (*see* Figure 1.3[1]). This PSS shift is known as *the recalibration effect.*

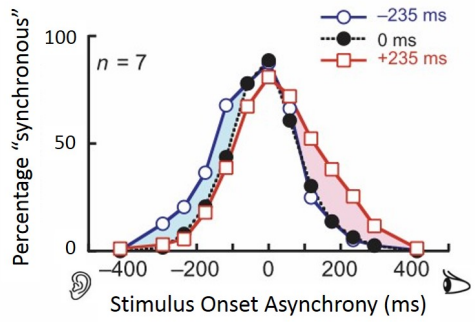

Figure 1.3: Proportion of synchronous responses as a function of SOA and the SOA that was presented prior to testing. When participants were previously exposed to auditory-leading (blue) stimuli, they were more likely to perceive auditory-leading stimuli to be synchronous during testing, and vice versa for visual-leading (red) asynchronies [38].

Vroomen et al. [40], investigated the recalibration effect, and exposed participants to the following SOAs during the adaptation period (-200ms, -100ms [auditory leading], 0ms [synchronous], 100ms, 200ms [visual leading]). As

[1]In Figure 1.3, it appears that the PSS shifts are manifested as a widening of the TBW in the direction of the exposure asynchrony. However, the authors did not report that the TBW was significantly affected by exposure to the asynchronous stimuli.

the SOA used in the adaptation period increased, the recalibration effect did not proportionally increase. The authors reported that there were no recalibration effects for adaptation asynchronies larger than 350ms in a separate experiment. It was hypothesised that recalibration does not take place after large asynchronous lags because the cross-modal stimuli are not combined during the adaptation phase [40].

Temporal recalibration does not only occur with simple stimuli such as flashes and tones. Navarra et al. [41], presented participants with videos of speech (or a musical pattern played on a piano) with an auditory lag of 300ms. Participants performed a TOJ task on asynchronous flashes and tones that were presented throughout the videos. In contrast to Fujisaki et al. [38] and Vroomen et al. [40], the PSS did not shift in the direction of the adapted stimulus. However, exposure to the asynchronous speech increased the length of the TBW. Therefore, when presented with asynchronous speech, the test stimuli had to be more temporally separated for participants to be able to accurately judge their temporal order. As in Vroomen et al. [40], the widening effect was not present when the adaptation stimulus was very large. The speech and music adaptation stimuli affected perceptions about the temporal order of beeps and flashes, leading to the opinion that recalibration effects may not be stimulus specific. Finally, when the test and adaptation stimuli are both speech streams the PSS did shift in the direction of the exposed asynchrony as expected [42, 43].

It has been demonstrated that temporal recalibration is not specific to audition and vision, but is also present in auditory-haptic and visual-haptic integration [44, 45]. Recalibration to asynchronous sound/light can modulate detection speeds and response times [46]. However, a parametric analysis of recalibration in sound/light/touch pairings did not detect PSS shifts in sound-touch stimuli or light-touch stimuli after adaptation to the same pairing [47]. Harrar et al. [47] additionally recorded response times to unimodal stimuli after adaptation to each sensory pairing; their findings were not in agreement with Navarra et al. [46] with only detection speeds of visual stimuli being modulated. The generality of temporal adaptation mechanisms between multiple modalities is not known.

There is disagreement over the underlying mechanisms of recalibration. Some authors have hypothesised that adaptation influences the speed at which auditory information is processed [46]. Others propose that the timing mechanism for auditory-visual processing is separate from those involving touch [47], and a single multisensory mechanism has also been hypothesised [44].

More recently Van der Burg et al. [39] demonstrated that recalibration effects can be elicited after a single exposure to asynchrony, this is in contrast to previous work, which required sustained exposure for adaptation to occur. In these experiments, [39] participants judged whether transient beeps and flashes (presented at a range of stimulus onset asynchronies: ±0ms, ±64ms, ±128ms, ±256ms, and ±512ms) were synchronous or not. The experimental session was extended to 120 trials per asynchrony (1200 trials in total), in order to use the stimulus onset asynchrony (SOA) of the previous trial as a model factor in their analyses. Their analysis showed a main effect of trial $t-1$ SOA on the mean PSS of trial t, and group differences in mean PSS for trial t when preceded by either auditory leading or visual leading stimuli. The rapid recalibration was strongly asymmetrical, with visual-first $t-1$ asynchronies more strongly affecting PSS on trial t than auditory-first asynchronies (*see* Figure 1.4). The asymmetry of the effect was attributed to the natural delays in auditory-visual stimuli due to the faster transmission of light than sound. It was additionally shown that the rapid recalibration effect occurred regardless of whether the previous trial was perceived to be synchronous, and whether or not the previous trial required any conscious decision on synchrony. The authors suggested that the rapid recalibration mechanism is a low-level sensory effect that maximises benefits of auditory-visual integration, rather than a top-down process, whereas previous investigations of adaptation documented a top-down process that linked correlated events through sustained exposure.

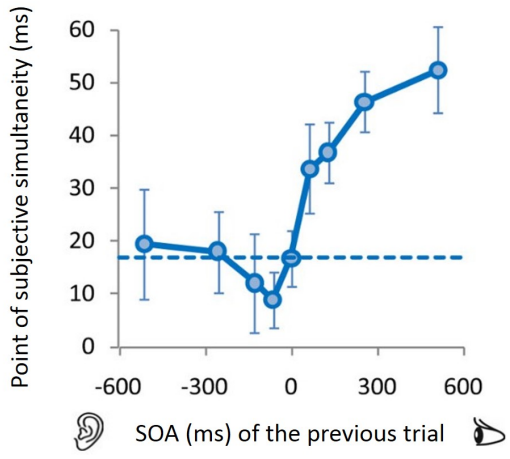

Figure 1.4: The mean PSS on trial t as a function of the SOA of trial $t-1$. The dotted line indicates the average PSS for an SOA of 0ms on trial $t-1$. The SOA of one trial can affect the percieved simultaneity of subsequent trials. The rapid recalibration effect occurred for large visual-leading asynchronies, but only for small auditory-leading asynchronies. Figure adapted from [39].

Further investigations of rapid recalibration for transient stimuli have shown that it is necessarily an auditory-visual process: it does not occur for auditory-haptic and visual-haptic stimuli [48], and does not occur for unimodal stimuli [49]. Rapid recalibration has been demonstrated using using more ecologically relevant stimuli (brief video clips in which a speaker pronounced a single syllable with various SOAs). A first experiment showed that the point of subjective simultaneity was highly dependent on the modality order of the previous trial. In a second experiment, the speaker in the video was changed at random between trials, but rapid recalibration continued to occur. In a final experiment, male speakers' syllable pronunciation was overlaid with a female voice and vice versa. This experiment violated the unity assumption of multisensory integration by introducing semantically incongruent auditory-visual stimuli, but rapid recalibration consistently occurred [50]. These findings promote the hypothesis that rapid recalibration is a low-level mechanism driven by basic temporal factors.

Similarities should be noted here between the rapid recalibration effect and the *serial dependency effect*. Serial dependence occurs when recently perceived or remembered information affects current information processing [51]. For example, in [52] participants were presented in each trial with a Gabor patch, then 1s Gaussian noise, a 250ms fixation, and finally an approximate 1100ms response bar. Participants were asked to judge whether the response bar was oriented more clockwise or anti-clockwise that the Gabor patch. It was shown that the error on the current trial was positively biased in the direction of the Gabor patch presented in the previous trial, and significant effects were present when accounting for the second and third previous trials. However, the rapid recalibration effect differs significantly from the serial dependence effect because of the clear asymmetry in it's effect size.

Rapid recalibration paradigms show PSS shifts after a single exposure, but it is not known whether mechanisms that underlie this process are the same as in classical recalibration studies in which there were extended adaptation periods. To further investigate the differences between rapid and classical recalibration, [53] used a combination of experimental paradigms, using an adaptive period and examining trial $t-1$ order effects, to investigate timescales of each phenomenon, and determine whether they operate independently. Their results confirmed a dissociation of effects between classical and transient recalibration. Sustained adaptation over 3 minutes caused large PSS shifts that decayed after around 1 minute, whereas rapid recalibration effects were consistent throughout the experiment. The authors proposed that rapid recalibration is mediated by shifts low-level temporal alignment between modalities, and that prolonged exposure to asynchronous stimuli encourages top-down feedback which enacts longer term change [54]. Clas-

sical and rapid recalibration were recently confirmed to operate on different timescales [55].

It is important to make a distinction between temporal recalibration and temporal adaptation. Temporal recalibration refers to a shift in the measured state of the mechanism responsible for determining synchrony *after* adaptation has occurred. I define temporal adaptation to be an increase in perceptual synchrony over time, as a result of ongoing exposure to asynchronous stimuli. A key difficulty in the interpretation of temporal recalibration studies is that recalibration, the *effect* of adaptation, can be measured, but adaptation cannot. For both classical and rapid recalibration studies, it is not known whether the exposure to asynchrony alters low-level mechanisms for judging simultaneity, or simply alters high-level subjective criteria. More research is needed to clarify this distinction.

Stimulus type, feedback training, and adaptation exemplify the extent to which the the perceived simultaneity of events can be modulated by external factors. This may be reflective of malleability in underlying neural processes. For example, emerging evidence supports the notion of brain as an "active" sensor, in which momentary context is used to generate predictors for future events [56–59]. Neural oscillations are central to active sensing, and it has been argued that the brain encodes information about the temporal structure of stimuli using oscillatory mechanisms. Unconscious temporal predictions about stimulus onset are then made unconsciously via the likelihood of stimulus occurrence during a period of high excitability [60–63].

A common characteristic in many environmental stimuli is temporal regularity [62]. For example, speech has amplitude modulations with regular frequencies between 2 and 7 Hz [64, 65]. It has long been posited that the brain makes use of these natural temporal structures in order to manage attentional resources [66, 67].

Dynamic Attending Theory (DAT) is a framework within which I can describe how attention can vary over time. Three key properties of this framework are: 1) attention levels oscillate at an internal frequency when there is no external stimulation; 2) attention can become coupled with external, naturally rhythmic stimuli through entrainment; 3) the oscillation of attention will return to the internal frequency after external stimulations are removed. A specific prediction of DAT is that the entrainment of attention to external stimuli becomes more accurate over time, and that perception is improved during attentional peaks [68](*See* Figure 1.5). There are a wealth of studies which provide evidence in support of these predictions [68–76].

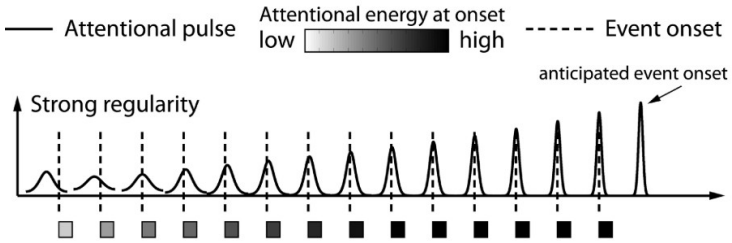

Figure 1.5: DAT predicts that attentional oscillations become entrained to external stimuli, and that this entrainment is strongest when the external stimulus is regularly timed. Here, attentional pulses, or periods in which attention is focused, are shown during a during a stream of regularly timed events. Over time, epochs of high attention become progressively shorter, and the amount of attentional energy used during each epoch is heightened. During these periods of heightened attention, the reliability of stimulus perception is theoretically increased. Figure adapted from [67].

For example, Large et al. [68] presented pairs of consecutive tones, and asked participants to make judgements about the delay between their onset. Participants first heard the standard interval, then a comparison interval, and judged whether the comparison interval *same* or *different* to the standard. Before the presentation of the standard interval, participants were passively exposed to regular, isochronous tones. The authors showed that the relationship between the duration of the standard interval and the inter-stimulus interval (ISI) of the preceding cues significantly affected the accuracy of participants' comparisons. When the standard interval was the same duration as the preceding ISI, the accuracy of participants' judgements increased. Temporal discrimination decreased as the standard interval deviated from the interval of the preceding train of tones. In the context of DAT this provides evidence that attention is being entrained to the preceding cue stimuli, because judgements about the comparison interval were most accurate when entrainment was upheld by the standard interval.

Another study of DAT showed that temporal patterns and regularity can improve visual perception [70]. In this experiment, participants were required to discriminate between left and right oriented 45 ° visual gratings (Gabor patches) that were overlaid with Gaussian noise. Participants simply responded *left* and *right* to the gratings, and discrimination performance increased as noise level decreased. Importantly, the target stimuli were presented within a stream of distractor stimuli; these were visual Gaussian noise stimuli that were presented in either a regular or irregularly timed sequence. Participants made more accurate judgements about orientation when target stimuli were presented in the regularly timed sequence. This is important because it suggests that entrainment to external stimuli does not only affect temporal perception, but can improve sensory sensitivity across senses.

Recently, evidence has begun to emerge suggesting that temporally regular stimuli can influence perception across senses. For example, Escoffier et al. asked participants to judge whether a picture of a face or a house was upright or upside-down, and timed their responses [74]. Three conditions determined the temporal context of these judgements. Firstly, the pictures were presented alone, with no surrounding temporal structure. Secondly, the pictures were presented alongside an accompanying musical beat. Importantly, this condition was timed such that the target presentation occurred on-beat[2]. Finally, the pictures were presented off-beat. There was a significant effect of temporal context on the response times of judgements. Response times were slowest when there was no accompanying rhythm, and fastest when the target stimuli were presented on-beat. This supports the view that attentional entrainment is not specific to any individual sense. Via DAT, it is clear that both temporal discriminability and multisensory processing can be affected by temporal regularity. However, it is not clear whether dynamic attention directly affects the perceptual synchrony of multisensory events.

Temporal context may affect perception via its effect upon attention, but there may be more direct physiological effects of temporal context. *Active sensing* describes how temporal fluctuations in sensory performance may result from slow neural oscillations that can be entrained by external stimulation. For example, corticocortical oscillations may be tied to epochs of increased excitatory activity in, say the primary auditory cortex, which may affect sensory processing during the epoch [59–62]. The phase and frequency of these neural oscillations could become entrained to external stimuli via phase resetting mechanisms [77–82], or could result from the intrinsic frequency of motor sampling routines [62, 63]. It is likely that, the true contribution of temporal context to perception combines ideas from both frameworks. Indeed recent reviews have highlighted the similarites between the two [67].

1.2.2 Temporal Congruence

Detection of temporal incongruence may be achieved by the moment to moment evaluation of the synchrony of sensory components. However, it is important to consider that natural stimuli are rarely transient, and often contain rich temporal structure [64, 83]. When presented with streams of auditory and visual stimuli, we are able to match temporal structures over time, and matching performance is greatest when the temporal structure of

[2]To avoid an auditory confound between the beat and no beat conditions, the authors presented the target stimuli where the beat was *predicted*, but without any sound.

the streams are most complex [84]. It is likely that such matching techniques are used to detect temporal congruency, especially when stimuli are degraded. For example, the detection of spoken sentences in noise is greatly increased for temporally congruent stimuli compared with temporally incongruent stimuli [13].

In a recent study, Crosse et al recorded electroencephalography (EEG) while participants heard and/or viewed spoken conversation [10]. The auditory component of the stimulus was presented with or without degradation (noise mask). As expected, speech intelligibility was improved when participants were presented with auditory and visual components compared with when either modality was presented in isolation. Compared against the unisensory condition, in multisensory speech there was a greater increase in intelligibility when the auditory component was degraded than when it was not degraded. This demonstrates the principle of inverse effectiveness. The authors then used the EEG recordings to reconstruct the speech envelope using a decoding technique with which data over a time range could be incorporated into the model. The time window was increased incrementally from 0ms (instantaneous EEG data) to 500ms. When speech was degraded, the decoder required a larger data range to accurately reconstruct the speech envelope. The authors argued that inverse effectiveness is reliant on the integration of information over large temporal windows.

1.3 Spatial and Semantic Factors

Spatial coincidence and semantic consistency are crucial factors in multisensory integration. It is important to understand their contribution to integration to give further context to the investigations presented in this thesis.

1.3.1 Spatial Congruence

Spatial congruency is determined by the source locations of the stimuli. If unisensory components of a multisensory stimulus originate in the same locations, the multisensory stimuli are said to be spatially congruent. It has been shown that the spatial orienting of attention occurs more quickly and accurately for multisensory signals than for unisensory signals [85, 86] (for a review, *see* [87]). When signals are spatially congruent, multisensory processes can confer a perceptual benefit. For example, Frassinetti et al. [6] varied the spatial proximity of an auditory transient tone and a light, and

measured the detection rate of the target light. The detection rate was greatly improved when the sound and light were spatially congruent.

By manipulating spatial congruence whilst maintaining temporal and semantic congruence, it is possible to trick the brain into perceiving the signals as originating from a different source: this is known as the ventriloquism effect [88]. In this multisensory illusion, the relative certainty of spatial location provided by visual signals overrides any spatial disparity with respect to auditory signals, and sounds are perceived to originate from a different location then their true source [89]. It should be noted that auditory information can also influence the perceived location of a visual stimulus, but only when the location of the visual stimulus is made less reliable [90]. This had led to the assertion that spatial information from both senses are combined optimally to reveal most about the stimulus, regardless of modality.

1.3.2 Semantic Congruency

The estimation of semantic congruence requires the comparison of a multisensory signal with learned associations of multisensory stimuli. For example, a "meow" sound is semantically congruent if it appears to originate from a cat, but it would be incongruent if it originated from a dog. Semantic congruence can also refer to learned meanings and definitions (e.g. the word "red" and the colour blue are semantically incongruent). It has been demonstrated that reaction times to semantically congruent auditory-visual stimuli are decreased relative to unimodal stimuli, and that reaction times are increased for semantically incongruent stimuli [20]. More recently it has been shown that target detection and identification can also be improved when congruent auditory-visual stimuli are presented, and diminished when the stimuli are incongruent [19, 21].

1.4 Neural Correlates of Multisensory Integration

I have reviewed the temporal, spatial, and semantic factors that influence multisensory behavioural phenomena. These stimulus characteristics are often manipulated to elicit multisensory responses, in order to understand the neural basis of these phenomena. In early research, the primary goal of such investigations has been to identify multisensory regions and examine their

properties. More recently, significant insights have been gained by investigating how these regions functionally interact. In the following sections, I will provide an overview of multisensory integration in the brain.

1.4.1 Integration in Sensory Cortices

Sensory perception is traditionally investigated one modality at a time. However, there is evidence to suggest that multisensory processes can affect sensory processing at an early level, within early sensory cortices. For example, Calvert et al. [91] presented participants with auditory-only, visual-only, and auditory-visual speech, and recorded neural activity with fMRI. When auditory-visual speech was contrasted against the sum of unimodal conditions, there was significant activation in the primary auditory cortex and visual area V5. This is analogous to multisensory enhancement.

A recent study combined MEG with fMRI to give high spatiotemporal accuracy and presented participants with auditory-only, visual-only, and auditory-visual combinations of tones and a black and white checkerboard pattern. Initial responses to multisensory stimuli occurred in the primary auditory cortex after 20ms, and auditory-driven activity was present in the primary visual cortex within 10ms. Initial activation in the visual cortex occurred after 40ms, and visual-driven activity was present in the auditory cortex within the next 55ms [92]. It is possible that such low latencies are facilitated by afferent connections between the sensory cortices. Diffusion tensor imaging (DTI) studies have shown that white matter tracts connect auditory and visual cortices directly [93, 94].

A number of recent studies have demonstrated that the phase of oscillations in the visual cortex can be "reset" by auditory stimuli [77, 80, 82]. For example, Mercier et al. [80] presented participants with sequences of simple auditory-only, visual-only, and auditory-visual stimuli with a random ISI between 750ms and 3000ms and made EEG recordings. Auditory-only stimuli reset the phase of oscillations, and in some cases elicited event-related potentials in the visual cortex. Phase resetting also occurs in the auditory cortex as a result of visual stimulation [56, 95] (for a review *see* [96]).

Phase resetting is an important multisensory mechanism. A recent investigation has demonstrated that the relative timing of events between sensory modalities is encoded in the phase of neural oscillations [97]. In this study, participants were presented with a stream of transient flashes and tones with an SOA which switched between 0ms (4 repetitions), -200ms (8 repetitions),

and 200m (8 repetitions). The auditory-visual stimuli were presented with a regular ISI to elicit robust frequency tagging of cortical responses. EEG recordings were taken throughout the experiment. There were systematic shifts in the phase of oscillations in the auditory cortex, which matched the asynchrony of the presented stimuli.

1.4.2 Multisensory brain regions

The first brain region that was shown to demonstrate multisensory properties was the superior colliculus (SC). Meredith & Stein [98] demonstrated that many neurons in the deep laminae of the cat SC were highly active upon presentation of an auditory-visual stimulus, but much less active when auditory or visual stimuli were presented in isolation [98]. This seminal paper led to increased research interest in multisensory integration, and subsequent studies revealed many important properties of multisensory neurons. These include response enhancement, inverse effectiveness, and response depression [17, 99–102].

Many multisensory brain regions have been identified using imaging techniques. Calvert et al. [14] carried out one of the first studies of this type. In this study, participants viewed an alternating black and white checkerboard pattern and heard 100ms white noise bursts. The white noise bursts were presented in time with the checkerboard alternation, to produce temporal congruency, or with a random offset, to produce temporal incongruency. The auditory and visual stimuli were also presented in isolation. Criteria were adapted from early electrophysiological studies to identify multisensory regions. The authors contrasted the multisensory response during temporal congruence against the sum of the unisensory responses to identify regions exhibiting multisensory enhancement in the BOLD effect. In addition, the multisensory response during temporal incongruence was contrasted against the maximum unisensory response, to identify regions that exhibit multisensory depression. Their search yielded an area in the left superior temporal sulcus (STS) that was especially sensitive to multisensory stimuli. Evidence from multiple studies now support that the STS is a multisensory region [14, 16, 103–108].

Other multisensory brain regions include the posterior parietal cortex (PPC) [109], and prefrontal cortex (PFC) [110, 111]. Though the exact mechanisms that govern processes in the SC, PPC, and PFC remain to be determined, many studies have explicated the multisensory function of these areas. Saccadic eye and head movements are controlled by the SC [112], this involves

coordination between auditory and visual frames of reference. Saccadic response times are faster when an auditory and a visual stimulus originate within the same receptive field [113]. The PPC is involved in goal directed limb movements, which requires correspondence between sensory receptive fields [114]. In addition, there is evidence that the PPC maintains correspondence between spatial frames of reference between senses [115, 116]. There is evidence that the PFC is involved in integrating auditory-visual communication signals [111], and categorising semantic information [117].

The function of the superior temporal sulcus (STS) in relation to multisensory processes is less clear. High definition fMRI has also revealed discrete patches within the STS that respond to either auditory, visual, or auditory-visual stimuli [104]. In some studies the STS has a larger response to auditory-visual speech stimuli than to other types of stimuli [16, 118], and is more active during object recognition [105, 108]. This suggests that the STS has a role in semantic categorisation. However, similar results can be found in this area with auditory-visual stimuli that have no semantic or lexical content [14, 119]. This has led to the view of a more general multisensory role for the STS [2].

Several brain regions have been associated with temporally congruent and incongruent auditory-visual stimuli. Typically, activity in these areas is increased for congruent stimuli and decreased for incongruent stimuli, but this is not always the case. The superior colliculus is maximally responsive to auditory-visual stimuli that are presented simultaneously, and is significantly less responsive to unimodal stimuli or less temporally proximate stimuli [15]. More recently, fMRI studies have associated cortical areas with temporal congruence, including the inferior/superior frontal gyrus, superior temporal sulcus, and inferior/superior parietal lobule [11, 14, 16, 91, 104, 105, 108, 119–121]. The association of these areas with temporal congruence specifically, rather than semantic or spatial congruence, is debatable. However, overall brain activity patterns remain consistent between studies with varying stimulus types, e.g. speech [16, 122], transient beeps and flashes [14, 107], and speech-like ellipses and tones [119], suggesting that these regions are responsible with for low-level temporal correspondence calculation, regardless of their involvement with higher level semantic consistency.

1.4.3 Connectivity

Multisensory integration involves functional connectivity between brain regions, at least between primary sensory cortices. Connectivity has been

investigated in several studies using psychophysiological interaction analyses or dynamical causal modelling [28, 29, 107, 122–125]. Some studies have reported connectivity between the superior temporal sulcus and primary sensory regions, whereas others showed connectivity between superior temporal sulci and inferior frontal gyri/sulci. Different stimulus types may account for these findings, for example Noesselt et al. [107] used sound bursts and transient colour changes in either temporally congruous or incongruous patterns to investigate temporal correspondence devoid of semantic content. This indicated that connectivity between the STS and primary sensory cortices is increased significantly during temporal congruence. Lewis & Noppeney [28] also found greater connectivity between primary sensory regions when using simple non-semantic sounds and shapes (a click sound accompanied a rotational movement). In contrast, stimuli used in [124] were auditory-visual video clips of tool use or musical instruments, subject to varying degrees of degradation, and in [122] participants were presented with temporally asynchronous spoken sentences. Both studies used stimuli with increased semantic content and reported increased connectivity between the superior temporal sulcus and the inferior frontal sulcus.

The majority of investigations of temporal congruence used either transient stimuli or stimuli with associated semantics. Laing et al. [126] examined neural activity and connectivity elicited by continuous stimuli that were devoid of semantic information, but retained key aspects of the temporal correspondences present in speech. Using amplitude-modulated tones and size-modulated shapes, the authors manipulated the modulation rate of either the auditory or visual stimulus to create temporally congruent and incongruent conditions. Activation in the superior temporal gyrus, precuneus, and intraparietal sulcus was increased when the stimuli were congruent, relative to when they were incongruent. Connectivity between the superior temporal gyrus and the inferior, middle, and superior frontal gyri was stronger for temporally incongruent stimuli compared with congruent stimuli. These results are consistent with Noesselt et al. [122] who presented speech stimuli, and manipulated delays between auditory and visual streams so that participants were likely to switch between synchronous, and asynchronous percepts. Crucially, in [122] connectivity between the STS and frontal regions was increased during the perception of asynchrony. Collectively, this suggests that effective connectivity between frontal regions and the superior temporal sulcus is important for monitoring temporal correspondence during physical asynchrony, even when semantic information is absent.

1.4.4 Multisensory integration and neuronal inhibition

Inhibition is an important mechanism of integration. For example, when auditory-visual signals are not aligned in space and time, multisensory processes inhibit the neural response [8, 15, 16, 91, 98, 100, 101, 127] (for a review *see* [2]). This is known as response depression.

The inhibitory neurotransmitter gamma-aminobutyric acid (GABA) is prevalent in the brain. Early experiments have used GABA antagonists to reduce the spatial selectivity of visual neurons [128] and the pitch selectivity of auditory neurons [129]. This suggests that GABAergic inhibition is important for sensory tuning. Until recent years it has not been possible to measure GABAergic inhibition non-invasively, however the development of spectral editing techniques has led to a significant increase in signal to noise ratio for in vivo measurements of GABA concentration using magnetic resonance spectroscopy (MRS) [130–132].

Although the measurement of GABA concentration is not a direct measure of GABAergic inhibition, investigators have discovered a plethora of correlations between GABA concentration and their research focus: GABA concentration in the frontal and parietal lobes declines with age [133], GABA concentration in the temporal lobe is decreased in patients with tinnitus and presbycusis [134, 135], GABA concentration has been associated with developmental disorders [136], and finally, GABA concentration is correlated with dosage of antipsychotics and anticholinergics in patients with chronic schizophrenia [137]. It has also been demonstrated the production of neural oscillations in the gamma frequencies is affected by the concentration of GABA [138–141]. It is thought that GABAergic inhibition causes increased power of neuronal oscillations, especially in the gamma band, via increased excitation of inhibitory GABAergic interneurons.

With respect to sensory processing, increased GABA concentration in individuals is correlated with a number of behavioural metrics, including orientation discrimination performance [142], tactile discrimination performance [143], tendency towards visual motion assimilation [144], and percept duration in visual bistability [145]. Edden et al. [142] measured orientation discrimination thresholds for sequential black and white circular gratings presented either vertically or obliquely, and magnetoencephalography (MEG) to measure stimulus induced peak gamma frequency and amplitude in the visual cortex when participants viewed square-wave gratings over regular periods. GABA-edited spectra were then acquired in the occipital lobe of each participant. The authors showed that subjects' orientation discrimination thresholds for

oblique stimuli were significantly negatively correlated both with gamma frequency and cortical GABA concentration. GABA was also significantly positively correlated with gamma frequency. The authors suggested that the synchrony of the neuronal assemblies required for orientation discrimination are more easily maintained at higher gamma frequencies, thereby increasing accuracy, and that these higher frequencies are facilitated by increased GABA concentration. Oscillatory maintenance and correction in neuronal populations via phase reset mechanisms are known to occur in multisensory processing, particularly in primary sensory cortices. If GABA concentration does influence unisensory processing in this manner, a correlation is likely to exist between GABA and sensory thresholds in a multisensory task.

To date, only one study has demonstrated a direct link between multisensory integration, cortical GABA concentration, and gamma band oscilliations. Balz et al. [146] measured the participants' illusion rate in the sound induced flash illusion, wherein the presentation of one flash (transient dot) with two or more noise-bursts leads to the illusory percept of an additional flash/flashes. The illusion rate was compared against source localised gamma band oscillation power and GABA concentration in the left superior temporal sulcus. Significant correlations were obtained between all three measurements. This is an important finding that extends observations of GABA influence from unisensory to multisensory perception, and demonstrates that increased GABA in a multisensory brain region can increase gamma band power and predict multisensory perception.

1.5 Overview of Experiments

In the next four chapters, I report the key justifications, methods, results, and conclusions of four experiments. Each of these experiments are driven by theories and literature discussed thus far.

I have identified that the mechanisms which compare the relative timing of multisensory signals are likely to be context driven, and that the temporal orienting of attention to a specific epoch may be driven by stimulus regularity[68–70, 74]. However, it is unclear whether attentional orienting affects the perceptual simultaneity of multisensory events. In the first experiment, I aim to discover whether attentional modulation, driven by stimulus regularity, can widen or narrow the TBW.

There is speculation that rapid recalibration, in which there is a shift in

the PSS after a brief exposure to asynchrony, and classical recalibration, for which the PSS shifts after long periods of exposure, are reliant upon separate neural mechanisms. While there is evidence to suggest that these phenomena occur at different timescales [55], the key difference between these experiments is the adaptation phase, about which very little is known. In the second experiment, I investigate temporal adaptation directly using a novel paradigm to track perceptual responses to asynchrony over time. This paradigm will make use of continuous, speech-like stimuli, allowing me to study adaptation in an ecologically valid context. I aim to clarify the distinction between the mechanisms of rapid and classical recalibration by understanding the temporal context of these measurements.

Temporal adaptation may manifest in the brain as a shift in oscillatory phase in the auditory cortex over time, to account for stimulus asynchrony [97]. However, to my knowledge no brain regions have been specifically associated with temporal adaptation, and fMRI has not been used to adaptation directly. In the third experiment, I manipulate the state of temporal adaptation using a paradigm similar to that of Chapter 3 and measure brain activity indirectly using fMRI. I expect that activation is increased during temporal adaptation, and aim to identify brain regions that are associated with this phenomenon. Additionally, I aim to replicate recent work which has highlighted that parietal and frontal regions can be sensitive to temporal asynchrony and incongruence.

Inhibition is an important mechanism of multisensory integration and sensory tuning. A recent investigation has shown a correlation between resting GABA concentration in the STS, and the prevalence of an auditory-visual illusion. However, it is not clear whether the concentration of GABA leads to improved multisensory enhancement. In the final experiment, I measure the concentration of an inhibitory neurotransmitter in two putative multisensory regions. With a robust behavioural paradigm, I derive a metric of integration strength: this is the extent to which auditory perception is augmented by the presence of a visual object. I aim to demonstrate correlation between the concentration of GABA and integration strength.

Chapter 2

Dynamic Attention and the Temporal Binding Window

In subsection 1.2.1 I introduced the concept of the temporal binding window (TBW) and of dynamic attending theory (DAT). The extent to which mechanisms underlying perceptual simultaneity between senses are influenced by the temporal orienting of attention is unknown. In this chapter, a temporal order judgement (TOJ) task is adapted to investigate whether the TBW is affected by attentional priming to the onset of a multisensory stimulus. In these tasks, participants will judge the relative onset timing of auditory and visual stimuli. To understand the connection between attention and multisensory integration in brain and behaviour, regular or regularly timed priming sequences are presented to make the onset of the target stimuli more predictable.

2.1 Introduction

The temporal binding window (TBW) is a measure of tolerance to asynchrony between multisensory stimuli (*see* subsection 1.2.1). The TBW varies in duration within the population and between demographic groups [147]. Additionally, the TBW is dependent on stimulus type, experimental method [31], and on previously presented stimuli or multisensory asynchronies [38].

Temporal regularity is known to affect sensory processing, and this may occur via a number of possible mechanisms. For example, orientation discrimination is improved when presented during the onset of a beat in a regular

pattern [70], and temporal discrimination is improved when test stimuli are presented within a regular pattern [68]. In one study, participants' detection and discrimination of visual stimuli was improved when the stimuli were presented in time with an auditory rhythm [73].

We use dynamic attending theory (DAT) to describe how attention is modulated over time by both exogenous and endogenous rhythms. Sensory performance is increased during attentional peaks, as predicted by the theory (*see* subsection 1.2.1). Some early empirical support for DAT was based upon temporal discrimination studies [68, 69, 76]. This effect of attention on temporal discrimination is apparent in audition, but the effect of attention on temporal discrimination between senses is less clear.

Few studies have investigated the relationship between the TBW and attention. In one study, participants performed a SJ task in which the visual stimulus was a bouncing ball, and the auditory stimulus was an short click. By judging the speed of the bouncing ball as it approached the surface, participants were visually cued to the onset timing of the auditory click. The authors additionally tested participants' judgements about flash-click pairs presented at the same asynchronies. Temporal discrimination was improved for the bouncing ball stimulus relative to the flash-click pair [32]. Given the effect of attention on the TBW in this study, it is plausible that the TBW may be affected by fluctuations in attentional resources as influenced by temporal regularity in accordance with DAT.

In the active sensing framework, neural oscillations can become entrained to, or have their phase reset by, external stimuli. Mounting evidence suggests that ongoing oscillations affect sensory processing [77–82]. For example, one study asked participants to detect visual gratings during a 6 second period after the presentation of an auditory tone. The visual grating was presented alongside the tone, or at one of twelve time points separated by 500ms. Target detection performance was periodic, and locked to the onset of the initial auditory tone. In addition, periodicity was strongly influenced by the co-occurence of auditory and visual stimuli. The authors suggested that the initial tone reset the phase of ongoing oscillations and that periodicity in sensory performance was elicited by the increased likelihood of depolarisation periods of high excitability.

In the context of the TBW, neural oscillations may provide a low level temporal mechanism: auditory and visual stimuli are more likely to co-occur if they belong to a common oscillatory peak. Therefore, neural entrainment elicited via an external rhythm could induce tighter constraints on the perceived togetherness of the auditory and visual components.

In this chapter I will investigate whether the accuracy of temporal order judgements are affected by the regularity of stimuli preceding the test stimuli.

2.1.1 Aims

I hypothesised that:

- Cues to the onset of a multisensory pair reduce uncertainty about their perceived synchrony, and narrow the TBW.
- Regularly timed cues to the onset of a temporal order judgement further reduce this uncertainty.
- Regularly timed auditory cues narrow the TBW more than visual cues.

To address these questions we conducted three experiments in which we manipulated the existence and regularity of cues that preceded a flash-beep test pair.

2.2 Experiment 1

The primary aim of this experiment is to determine whether the temporal binding window is shorter when participants are attentionally cued to the onset of a temporal order judgement with regular cues compared with irregular cues. As a control condition, I also tested participants judgements about temporal order without any cues.

2.2.1 Methods

Participants

22 healthy adults (9 male; age: mean = 24.0, standard deviation = 4.1 years) took part in the study. Participants responded voluntarily to advertisements and were given a small monetary compensation for their time. Participants gave informed consent to the study, and to the use of their data. The study

was approved by the Newcastle University Ethical Committee and was run in accordance with the Declaration of Helsinki.

Stimuli

All stimuli were presented using MATLAB with the Psychophysics Toolbox extensions [148, 149]. The auditory tone consisted of a 400Hz sinusoidal pure tone with a duration of 33ms. Auditory stimuli were presented at a comfortable volume (70dB SPL) held constant across participants, and were delivered via headphones (Sennheiser HD380 Pro). The visual stimulus was a white circle presented for 33ms. The circle subtended a visual angle of 6.5° × 6.5° (150px × 150px) and was presented on a Dell™ UltraSharp™ 19" flat panel display (model 1907FP). For multisensory stimuli, the visual and auditory were presented together with some delay between their onset.

Design

To measure the TBW, participants made judgements about the temporal order of multisensory stimuli (tone-flash pairs). The tone-flash pair about which participants made temporal order judgements was preceded by 3 tones or flashes. There were 2 cue conditions (regular, irregular) tested within subjects, and two cue modalities (visual, auditory) tested between subjects. There were 9 Stimulus Onset Asynchronies (SOAs) between the tone and flash: -200ms, -150ms, -100ms, -50ms (audio-leading), 0ms (synchronous), 50ms, 100ms, 150ms (video-leading).

Procedure

In the regular cue condition, the three cues and the onset of the TOJ stimuli were isochronous. For irregular cues, the first and third cue onsets were the same as in the regular cue condition, but the onset of the second cue was manipulated such that a regular pattern could not be formed (*see* Figure 2.1. For each cue modality (auditory or visual), the onset of the same component of the TOJ matched the regular timing of the cue. The ISI between the cue stimuli was 0.6s or 0.8s and this was counter-balanced between trials.

The time course of each trial is outlined in Figure 2.1. Each trial began with a fixation cross for 1s, followed by a blank screen for 0.5s. Three cues were

presented either regularly or irregularly, in the form of light flashes or tone bursts. The tone flash pair began one ISI duration after the final cue. A green response dot was presented 0.6s after the offset of the test pair, indicating that a response was required from the participant. Participants pressed a key to indicate whether they thought the tone or flash occurred first. The response dot remained visible until participants gave their response. After that, the screen turned black for 0.5s before the onset of the next trial. There were 16 trials per SOA, leading to 144 different trials per cue condition.

The experiment was run in two 144 trial blocks with cue condition and ISI interleaved. Participants read instructions detailing how to respond to the stimuli, and were allowed 15 practice trials without feedback. Finally, participants were able to take a short self-timed break every 64 trials.

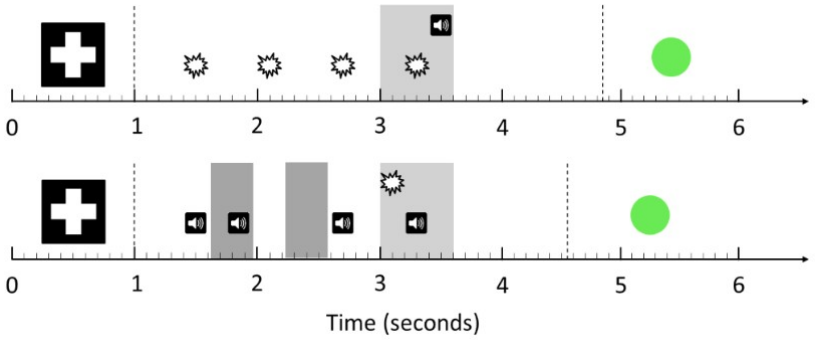

Figure 2.1: Experiment 1 trial time-course. In the regular condition (top), 3 isochronous cues preceded the onset of the TOJ stimuli. In the irregular condition (bottom), the onset of the second cue varied randomly (dark grey areas). This manipulation allowed the second cue to vary randomly, but not within 20% of the trial ISI from either the first and third cue, or from the original presentation time. The test pair were presented such that the stimulus with the same modality as the cue was presented one ISI after the third cue, and the stimulus with the opposite modality as the cue was presented with some SOA (light grey areas). Participants were asked whether the tone of flash occurred first, and responded via keypress when presented with the green dot.

In addition to the cued conditions, we tested temporal order judgements with no cue. Trials without cues were not interleaved with cue condition trials and were completed in a separate block prior to the cued conditions. In the no-cue condition, the onset of the test stimuli varied randomly. This was so that participants could not be cued to the onset of the stimuli via temporal regularity between trials. Each trial began with a fixation cross presented for 1 second After the fixation, the screen turned black. The tone-flash pair were presented with a random onset between 0.6s and 1.9s of the fixation offset.

Analysis

To estimate the TBW for each participant, we calculated mean flash-first response rates for each SOA in each cue condition. A cumulative Gaussian function curve was fitted to these means, and participants were removed from further analysis if the Gaussian provided a poor fit for the data ($R^2 < 0.5$). The TBW was calculated as the range of SOAs between which participants responded *"flash first"* 25% of the time, and the TBW when they responded *"flash first"* 75% of the time. For example, if the 25% threshold occurred at a -100ms auditory first SOA and the 75% threshold occurred at a 100ms visual first SOA, then the TBW would be 200ms. Thresholds were estimated using the corresponding value from the fitted Gaussian.

2.2.2 Results

Figure 2.2 shows the mean proportion of flash-first responses as a function of cue condition and SOA for $n = 22$ participants (11 for each cue modality). To investigate cue modality and regularity, we first performed a $9 \times 2 \times 2$ (SOA $\times$ cue regularity $\times$ cue modality) analysis of variance (ANOVA) on the proportion of flash-first responses with cue modality as a between-subjects factor and excluded the no-cue condition. There were no main effects present for cue regularity or cue modality and no interaction effects. There was a significant effect of SOA ($F(1.525, 30.508) = 35.568, p < 0.0005, \eta_p^2 = 0.640$).

After fitting the proportion of flash first responses, 9 participants were removed from further analysis of the TBW because the cumulative Gaussian function did not provide a good fit for the data ($R^2 < 0.5$; $n = 13$ remaining with 6 cued by visual stimuli, and 7 cued by auditory stimuli). This represents 41% of our cohort. The TBW for these participants is plotted in Figure 2.3. To assess cue condition and modality we performed a 2×2 ANOVA of cue regularity (regular, irregular) and cue modality (tone, flash). There was a significant effect of cue regularity on the TBW ($F(1, 11) = 5.986, P = 0.032, \eta_p^2 = 0.352$), which was shorted when the cue was regular (mean 84.8ms, standard error 10.9ms; collapsed over cue modality) than when the cue was irregular (mean 101.9ms, standard error 14.1ms; collapsed over cue modality). I performed post-hoc pairwise comparisons of the cued TBWs against a control condition with no cue, there was a significant difference between TBWs without a cue and TBWs with a regular cue after correction for multiple comparisons (regular cue vs. no cue: $t(12) = 2.627, p = 0.020$; irregular cue vs. no cue: $t(12) = 1.478, p = 0.165$).

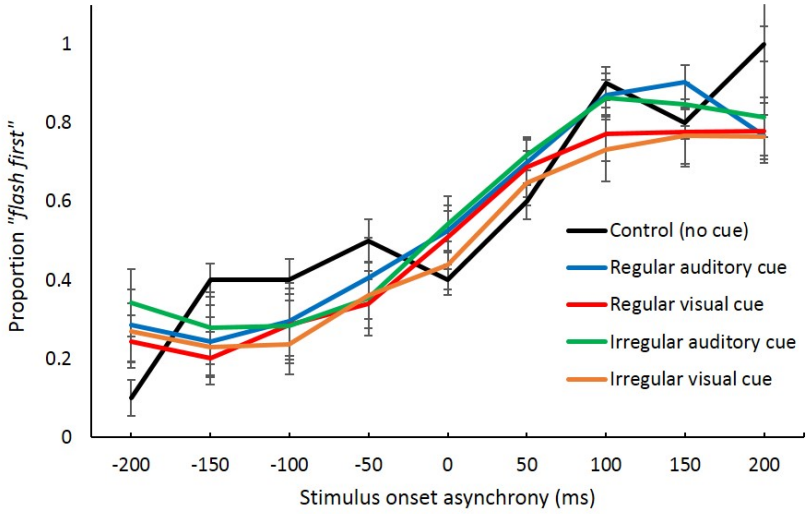

Figure 2.2: Flash first response rate for Experiment 1 plotted as a function of SOA, cue condition, and cue modality. There were 22 participants in the no-cue condition, and 11 participants for each cue modality. Error bars represent the standard errors of the means.

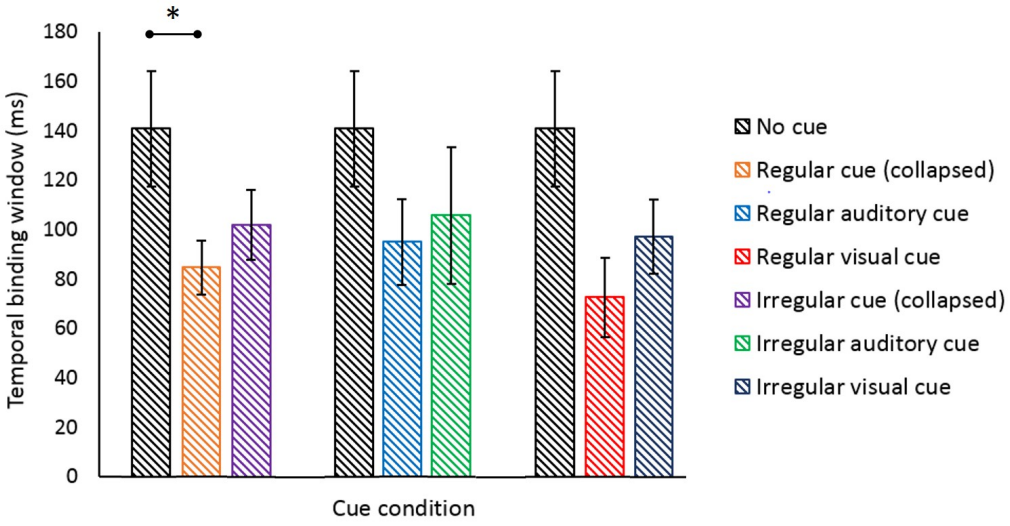

Figure 2.3: Mean TBWs by cue condition and cue modality. Error bars represent the standard error of the mean. There were $n = 13$ participants in the no-cue condition (black), $n = 6$ participants in the visual cue condition (red, dark blue), and $n = 7$ participants in the auditory cue condition (blue, green). Regular and irregular cue conditions that have been collapsed across cue modality are shown in orange and purple respectively. There was a significant decrease in the duration of the TBW in the regular cue condition compared to the no-cue baseline (collapsed over cue modality) as indicated by the asterisk.

2.3 Experiment 2

In section 2.2, we reported an effect of cue condition on the TBW in two separate analyses, but no effect of cue modality. To gain more certainty that these effects were caused by the rhythmic entrainment of attention, we carried out a related and more focused experiment. Firstly, we discarded the between-subject manipulation of cue modality in order to focus on cue regularity. We chose to use auditory cues only, because beat perception is poorer for visual rhythms than for auditory rhythms [150]. Additionally, we increased the number of cues preceding the test stimulus from 3 to 6. This bolstered differences between the isochronous and non-isochronous conditions, as well as reinforcing a rhythmic percept. Finally, we did not test a no-cue condition in order to focus on judgements of temporal order after regular and irregular cues. Together, the above manipulations provide a more focused investigation into the effects of dynamic attention on the temporal binding window.

2.3.1 Methods

Participants

8 healthy adults took part in the experiment. More detailed demographics are not available for these participants. Participants responded voluntarily to advertisements and were given a small monetary compensation for their time. All participants gave their informed consent to the procedure and to the use of their data. The experiment was approved by the Newcastle University Ethical Committee and wsa ran in accordance with the Declaration of Helsinki. These participants did not take part in any of the other experiments in this study.

Stimuli

Stimuli were as described in subsection 2.2.1.

Procedure

As in Experiment 1, participants made judgements about the temporal order of a tone-flash pair. Preceding the test stimuli were 6 tones in either a regular, or irregular pattern. Regular cues had an ISI of either 600ms or 800ms, counter-balanced across trials. The auditory component of the test stimulus was presented one ISI duration after the final cue to maximise temporal onset anticipation. The accompanying flash stimulus was presented at some SOA relative to the tone.

In the irregular cue condition, cues 2, 3, 4, and 5 were offset relative to the onset of the corresponding cue in the regular condition. The offset was a random value such that the new onset was not within 20% of the ISI of the preceding or subsequent cue, nor within 20% ISI of the original onset (*see* Figure 2.4).

There were 9 SOAs: -200ms, -150ms, -100ms, -50ms (auditory leading), 0ms (synchronous), 50ms, 100ms, 150ms, 200ms (visual first). There were 16 trials per SOA, leading to 288 trials in total. There were two blocks of 144 trials. Participants had a self timed break every 64 trials, and took a longer break between blocks. Finally, participants read the experiment instructions and carried out 15 practice trials without feedback before completing the experiment.

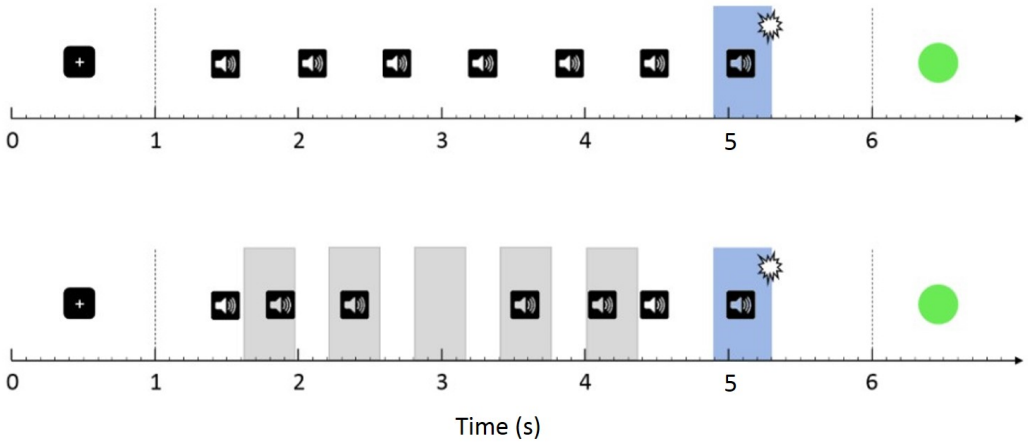

Figure 2.4: Regular and irregular cue trial time-courses for Experiment 2. A fixation cross was presented for 1 second. The first cue stimulus was presented 0.5s after the offset of the fixation cross. In regular cue condition trials (top), 5 further cues and finally the test stimulus were presented with an ISI of either 600ms or 800ms. In irregular trials (bottom), the onset of cues 2—5 varied randomly between the grey areas.

Analysis

Analyses were as described in subsection 2.2.1.

2.3.2 Results

Figure 2.5 shows the mean flash-first response rate for TOJs with 6 regular or irregular cues. We first performed a 9×2 (SOA $\times$ cue regularity) repeated-measures ANOVA on the proportion of flash-first responses. There was a significant effect of SOA ($F(1.520, 10.640) = 19.568, P < 0.0005, \eta_p^2 = 0.737$), but no effect of cue regularity (*see* Figure 2.5). There were also significant linear and cubic trends reflecting the response curve ($F(1, 7) = 25.527, P = 0.001, \eta_p^2 = 0.785$ and $F(1, 7) = 10.756, P = 0.013, \eta_p^2 = 0.606$ respectively).

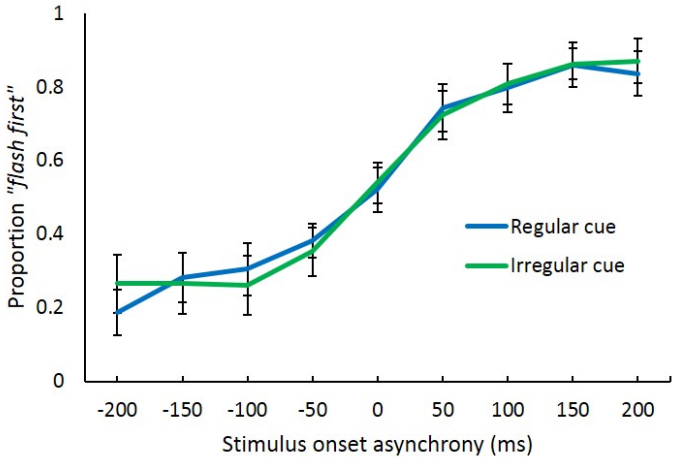

Figure 2.5: The mean flash first response rate for Experiment 2 with $n = 8$ participants, plotted as a function of SOA and cue condition. Error bars represent the standard errors of the means.

The TBW was calculated as described in subsection 2.2.1. One participant (12.5% of cohort) was removed from further analysis due to a poor fit of the psychometric function ($R^2 < 0.5$). The mean TBW for each condition is plotted in Figure 2.6. There were no significant effects of cue regularity on the TBW.

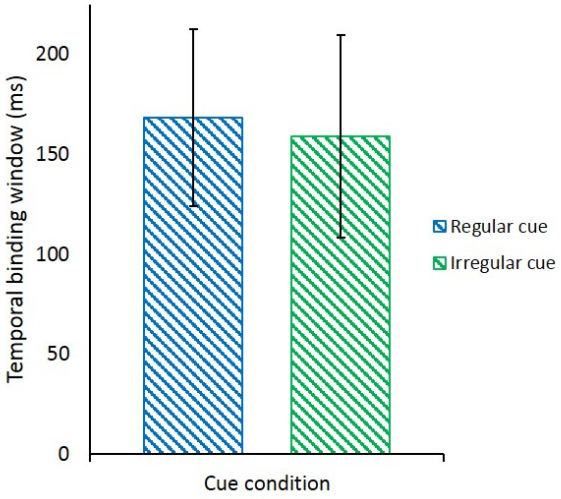

Figure 2.6: The mean TBW as a function of cue regularity for Experiment 2. Error bars represent the standard error of the means.

2.4 Experiment 3

In Experiment 2, I carried out an experimental procedure intended to focus on the effect of cue regularity. However, no effects or trends were present in the data. Additionally, cue regularity effects present in Experiment 1 were not replicated.

In the final experiment in this chapter, I altered the experimental paradigm used in Experiment 2 in three ways. First, to reduce experiment duration, I used an adaptive procedure to estimate TBW with fewer trials. This was to reduce fatigue as reported by participants in both Experiment 1 and Experiment 2. Second, I manipulated the auditory component of the test pair such that it was perceived to originate from the left or the right of the participant. This was to reduce the certainty about coupling between the TOJ test pair to increase reliance on preceding cues via inverse effectiveness. Finally, I reverted to 3 cues as in Experiment 1.

2.4.1 Methods

Participants

Twelve naïve participants (10 female; age: mean=20.3, standard deviation=2) took part in the experiment. All participants were healthy adults and did not take part in either Experiment 1 or Experiment 2. Participants responded voluntarily to advertisements and were given a small monetary compensation for their time. All participants gave their informed consent to the procedure and to the use of their data. The experiment was approved by the Newcastle University Ethical Committee and wsa ran in accordance with the Declaration of Helsinki.

Stimuli

All stimuli are as described in Experiment 1 and Experiment 2, with the exception of the auditory component of the test stimulus. This was a 400Hz tone 33ms in duration, but was presented more loudly in either the left or right ear to represent different stimulus locations.

Procedure

As in both Experiment 1 and Experiment 2, each trial consisted of a tone-flash pair and participants were asked to judge the order in which they occurred. Participants responded by pressing a key to indicate whether the flash or the tone occurred first. The intensity of the tone was increased in one ear such that the difference between left and right was 10 dB (average 70 dB). In addition, the phase of the sound wave was shifted by 90 degrees to the left or right. Thus, the auditory component of the test pair was presented more loudly in the left or right ear (counterbalanced between trials). The presentation of the test pair was preceded in all cases by 3 auditory cues (tones). Cue stimuli were presented equally loudly in each ear. Cues were regular or irregular, and their onsets were calculated as in Experiment 1: *Methods* (*see* Figure 2.1).

The four conditions were *regular-cue tone-first*, *regular-cue flash-first*, *irregular-cue tone-first*, and *irregular-cue flash-first*. Four interleaved QUEST [151, 152] staircase procedures were used to estimate the 75% correct thresholds for each condition (i.e. one staircase per condition). In both tone-first pro-

cedures, the test stimulus was always auditory-leading (0ms to -250ms). In the flash-first procedures, the test stimulus was always visual-leading (0ms to 250ms). The starting thresholds were 100ms and -100ms for flash-first and tone-first conditions respectively. Participants completed all trials in a single block and were allowed a self-timed break after 64 and 128 trials.

Analysis

Each QUEST procedure returns a probability distribution function (PDF) whose mean represents the calculated threshold. Participants were excluded from further analysis if the standard deviation of the PDF returned by the QUEST procedure was larger than 30ms, since the procedure could not accurately determine the threshold within 40 trials. The TBW was calculated as the difference between thresholds calculated in the tone-first and flash-first procedures[1]. The measurement of the TBW in this procedure is equivalent to Experiment 1 and Experiment 2, the only difference with Experiment 3 is that thresholds are estimated by the adaptative procedure instead of a psychometric fit.

2.4.2 Results

4 participants were removed from further analysis because the adaptive procedure did not converge (33% of cohort). The calculated TBWs are depicted in Figure 2.7. There was no significant difference between TBWs for the regular and irregular conditions ($t(7) = 0.334, P = 0.501$).

2.5 Discussion

In this study, I investigated the effect of regular and irregular cues on judgements of temporal order and the TBW. In the first experiment, I presented participants with auditory-visual stimulus pairs with a SOA. I preceded the test pairs with a regular or irregular auditory or visual cue. There was a significant effect of cue regularity on the TBW, and the TBW was shorter for regular cues than irregular cues. There was no effect of cue modality. In post-hoc tests, there was a significant increase in TBW duration when there

[1]Since they are a complement, the 75% threshold from tone-first procedures is equivalent to the 25% flash-first threshold.

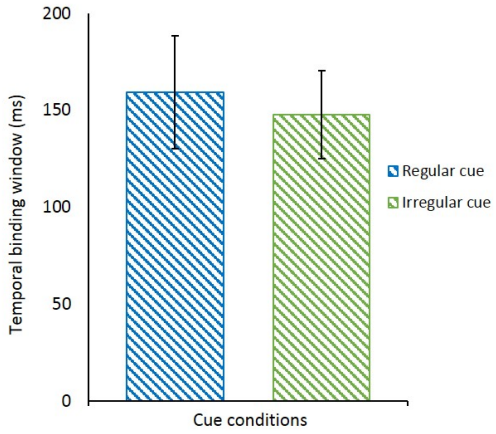

Figure 2.7: Mean TBW by cue regularity for Experiment 3. Error bars represent the standard error of the mean.

was no cue versus compared to when there was a regular cue. However, I could not repeat the cue regularity effect in experiments 2 or 3.

Temporal cues can increase sensory sensitivity and discriminability [153–156]. I predicted that the TBW would narrow when participants were cued to the onset of a TOJ. Evidence presented in subsection 2.2.2 suggests that discrimination performance in temporal order judgements is improved when subjects are cued to the onset of the test pair. This improved discrimination is equivalent to a shortening of the TBW.

Regular temporal patterns in the environment can contribute to further attentional tuning [68, 69, 71, 75](*see* subsection 1.2.1). I hypothesised that regularly timed cues could further reduce the TBW width. In Experiment 1, there was a significant effect of cue regularity on the TBW, which reflected shorter TBWs for regular cues then irregular cues. This evidence supports the hypothesis that stimulus regularity affects the mechanisms underlying perceptual simultaneity, however the result was not reproducible in further experimentation.

The TBW is known to be affected by a number of factors, including training [157], stimulus type [31–33], and experimental method [31, 32]. In [32], the TBW was narrower when participants made judgements about a bouncing ball stimulus, which provided a visual cue to the onset of an auditory tone. These results are in line with [32], and additionally show that cuing effects were not dependent on cue modality. It remains to be seen if there is an interaction between cuing and other modulatory factors.

Few studies have demonstrated a multisensory attentional effect. In [73],

participants made speeded judgements about simple stimuli whilst hearing a rhythm. Judgements were faster when stimuli were presented on-beat, rather than off-beat. In Experiment 1, participants were able to discriminate temporal order more accurately when judgements were made in time with an externally induced stimulus rhythm. This bolsters current evidence about the generality of DAT. In addition, this shows that attention can have an effect on how signals are matched temporally It is possible that increased attention to temporal synchrony reduces perceptual tolerances, thus reducing erroneous integration of mismatched stimuli. Additionally, the increased certainty associated with cued stimuli with low onset asynchronies could be used to increase gain during neural interactions, leading to superadditivity [15].

A key limitation of this work is present in the measurement of *"regularity"*. For example, the effects found in Experiment 1 may not have been reproducible because the irregular condition induced other regularities not accounted for by this manipulation. For example, throughout each experiment, the final cue stimulus and the component of the test pair with the same modality as the cue was always presented separated by one ISI. It is also possible that the regular cue condition did not reliably induce attentional entrainment. The inclusion of both visual and auditory cues in Experiment 1 is unlikely to have caused this discrepancy, since audition more reliably induces stimulus rhythm than vision [150]. To be certain of entrainment effects, future works should use EEG to monitor entrainment to the stimuli. This would constitute a quantitative verification of the effectiveness of the experimental manipulation.

Chapter 3

The Time-Course of Adaptation to Temporal Asynchrony

In subsection 1.2.1 I introduced the concept of recalibration to temporal asynchrony. There is speculation that rapid and classical recalibration rely on separate neural mechanisms. Classical recalibration relies on an extended period of exposure to asynchrony, known as the adaptation period. The adaptation period constitutes the main methodological difference between rapid and classical recalibration studies, yet auditory-visual temporal adaptation has, to my knowledge, not been measured directly. Here I introduce a new paradigm that uses continuous stimuli, and has temporal characteristics based on familiar speech patterns but contains no semantic content. This paradigm allowed me to measure the duration and strength of adaptation to a range of auditory-visual asynchronies. This gives insight into how the brain compensates for temporal asynchrony and the time-scales over which this compensation occurs.

3.1 Introduction

In auditory-visual recalibration studies, participants are asked to view auditory-visual stimuli with some temporal asynchrony, and in subsequent testing, are shown to have recalibrated their perception of synchrony [38, 40, 41] (*see* subsection 1.2.1). For example, if participants were exposed to a visual-leading asynchrony, they were more likely to perceive visual-leading asynchronies as

synchronous than if they had been exposed to no asynchrony, or an auditory-leading asynchrony. This is demonstrated as a shift in the PSS in the direction of the adaptation asynchrony. The key driver of temporal recalibration is the adaptation period.

In early papers, participants were exposed to the asynchrony for a number of minutes, and often re-exposed before each trial [38, 40]. More recently, rapid recalibration studies have also shown that recalibration can occur after a single presentation [39]. A key difference between results for rapid and classical recalibration is symmetry in the size of the PSS shift about the adaptation SOA. In rapid-recalibration studies, recalibration occurs more readily when the exposure stimulus is visual-leading. Whereas in sustained recalibration studies, effect sizes were typically similar for auditory-leading and visual-leading exposure stimuli. Is is not known whether temporal adaptation occurs equally for auditory-leading and visual-leading stimuli.

An investigation into the time-course of auditory-visual temporal adaptation is needed to better understand how perceptual simultaneity changes over time. For example, if adaptation can occur over a short duration, the extended exposure to asynchrony in classical recalibration papers was simply not necessary to elicit the effect. Additionally if adaptation occurs in the visual-leading direction only, then auditory-leading recalibration, as documented by classical recalibration experiments, may not have been driven by perceptual change.

To track the time course of adaptation, I presented participants with continuous, asynchronous auditory-visual stimuli and monitored perceptual simultaneity to track changes over time. By manipulating the auditory-visual delay of continuous stimuli during presentation, I observed real-time responses to stimuli that became asynchronous, and the following adaptation. This manipulation of synchrony during each trial is key to the measurement of adaptation, since the initial response to asynchrony can form a baseline measurement with respect to which I can measure the overall adaptation. The stimuli were based on speech, but were reduced to simple shapes and sounds. This maintains the ecological validity of the stimuli whilst removing semantic confounds.

3.1.1 Aims

Little is known about how the perceptual simultaneity of the adaptation stimulus changes over time. I recorded perceptual simultaneity throughout

the time course of adaptation, and hypothesised that:

- Auditory-visual temporal adaptation does not occur instantaneously.
- Auditory-visual temporal adaptation occurs to a larger extent for visual-leading stimuli than for auditory-leading stimuli.

3.2 Experiment 1

The primary aim of this experiment is to measure adaptation, the change in perceptual simultaneity over time, as participants hear and view temporally asynchronous stimuli. To do so, participants were asked to continuously report their perceived synchrony of an auditory-visual stimulus pair.

3.2.1 Methods

Participants

20 naïve, healthy participants took part in the study (mean age 20.25, range 9 years; 18 female). Participants responded voluntarily to advertisements and were given a small monetary compensation for their time. Participants gave written informed consent to the study, and to the use of their data. The study was approved by the Newcastle University Ethical Committee and was run in accordance with the Declaration of Helsinki.

Stimuli

Auditory and visual stimuli were initially derived from the IEEE sentences [158] which consist of 720 standardised sentences that are phonetically balanced. From this selection, 480 sentences were randomly selected without replacement. The sentences were then concatenated into 20 groups of 24, and pauses were added between sentences to mimic natural prose. The pause length was a randomly assigned value of either 300ms, 400ms, 500ms, 600ms, 700ms, 800ms, 900ms, or 1000ms and occurred between 50% of sentences. The groups of sentences were then full-wave rectified and low-pass filtered with a cut-off of 4 Hz. 50 second sections were selected from the resulting envelopes,

and used to define the amplitude of the auditory stimulus, and the size of the visual stimulus.

The visual stimuli consisted of size-modulated three-dimensional cuboids (as in [126]). The cuboid was created using 3D Studio Max 2016 (Autodesk, Inc.). To vary the size of the cuboid, the "spherify" modifier was applied to the centre ($1.0 \times 1.2 \times 4.0$ units [width $\times$ height $\times$ length]). The value of the modifier was allowed to vary between 0.1 and 0.5 (where 0 is no change and 1 creates a sphere) determined by the envelope created in the above procedure. The cuboid was rendered against a uniform black background from an oblique camera viewpoint. The bounding box of the cuboid subtended a visual angle of $13.7\,\text{deg} \times 13.7\,\text{deg}$ (300×300 pixels). The videos were saved as Quicktime movie files (3000 frames, 60 fps, H.264 compression) and were displayed on a DellTM UltraSharpTM 19" flat panel display (model no.: 1907FP) at a resolution of 1280×1024 pixels.

The auditory stimuli were pure tones with a carrier frequency of 250 Hz. The amplitude of the tone was determined by the envelope created in the above procedure. Changes in amplitude were then coincident with changes in the size of the visual stimulus, since they were created using the same envelope. In order to create stimuli that change from synchrony to asynchrony without any perceptible disruption, the auditory amplitude envelope was modified. For auditory-first asynchronies, a section was cut from the envelope at the required time point. For visual-first asynchronies, a randomly selected section of envelope was inserted at the required time point. The resulting concatenated envelopes contains one discontinuity for auditory-leading stimuli and two discontinuities for visual-leading stimuli. In all cases, discontinuities were compensated for using sine waves (*see* Figure 3.1). Audio files were created in MATLAB 2012 and were saved as stereo wav files with a 22.05 kHz sampling rate. Auditory stimuli were presented with Sennheiser HD380 pro headphones at 75 dB SPL .

Procedure

Participants were presented with videos comprising of an amplitude-modulated pure-tone and an animated shape. They were asked to continuously report whether or not the tone and shape were in synchrony with each other. To report their perception of the stimuli, participants held down the space bar on a computer keyboard if stimuli were synchronous, and released it when they appeared to be asynchronous. Responses were sampled 60 times per second. Each video began synchronously for 5s, and changed seamlessly to one of 6

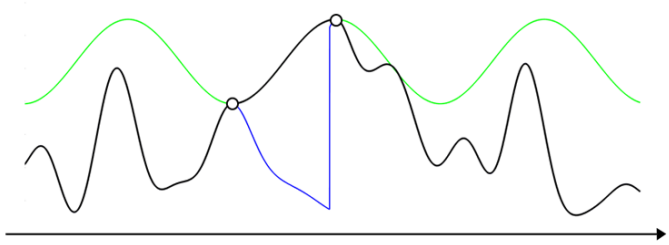

Figure 3.1: To manipulate the synchrony of the auditory and visual components of the stimulus, envelope discontinuities were patched using sinusoids. The section of envelope which has been cut and concatenated is shown in blue. A sine wave (green) was generated such that the stationary points of the wave are aligned with the nearest stationary points to the discontinuity. The resultant modified auditory amplitude envelope is shown in black.

asynchronies (-300ms, -200ms, -100ms [auditory-leading], 0ms [synchronous], 200ms , 300ms, 400ms [visual-leading], *see* Figure 3.2). The 100ms visual-leading SOA was not included as pilot testing indicated that participants were not able to distinguish that these stimuli differed from the synchronous condition. There were 12 trials for each asynchrony, leading to 84 trials lasting 50s each. To minimise fatigue, participants ran the experiment over two separate runs; each run contained an automatic break when half of the trials were completed.

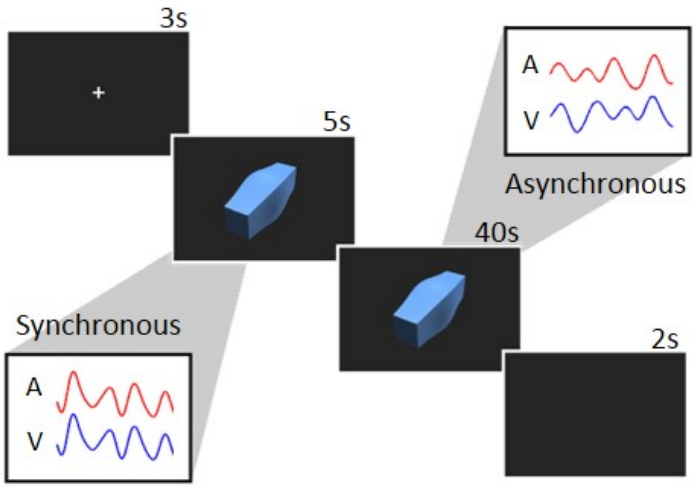

Figure 3.2: Experiment 1 trial time-course. Each trial began with with a fixation cross for 3s. The auditory-visual stimulus was then presented with 0ms SOA for the first 5s of each trial. The stimulus then transitioned seamlessly to the trial SOA. Participants continuously monitored the stimulus and reported their perceived simultaneity of the auditory and visual components. After stimulus presentation, the screen was blank for 3s before the onset of the next trial.

Analysis

Initially, trials in each condition were time-averaged. The resultant time-series were down-sampled to the average proportion synchronous during each second, yielding 45 data points per participant per condition. To determine where to begin fitting the data, a separate fitting procedure was used to estimate the point at which participants' responses were most impacted by the shift from synchrony to asynchrony. To do so, the following Gaussian function was fitted to the trial and time averaged data for 15s following the change to asynchrony: $f(x, a, b, c) = 1 - ae^{-\left(\frac{x-b}{c}\right)^2}$. Time series fits began at the minimum point of this initial fit, rounded to the nearest second (*see* Figure 3.3).

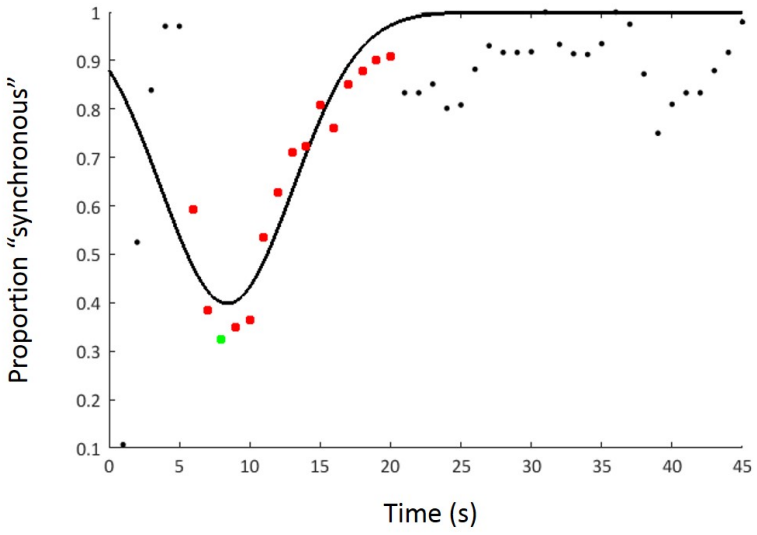

Figure 3.3: An initial fit is used to determine the starting data point of the final fit. An inverted Gaussian, shown in black, is fitted to the 15 datapoints that follow the stimulus transition (red). The start time for the final fit (green) is closest in time to the minimum of the fitted Gaussian

To model the responses to asynchrony, a scaled cumulative Gaussian was fitted to participants' responses from the pre-calculated start point until the end of the trial [159]. The fitting function was defined as

$$H(\alpha, \tau, \sigma, t) = \alpha G\left(\tau - \frac{\alpha}{2\sigma}, \frac{\alpha}{\sqrt{2\pi\sigma^2}}, t\right)$$

for $\sigma > 0$,

$$H(\alpha, \tau, \sigma, t) = 1 - (1 - \alpha) G\left(\tau + \frac{1-\alpha}{2\sigma}, \frac{1-\alpha}{\sqrt{2\pi\sigma^2}}, t\right)$$

for $\sigma < 0$, and $H(\alpha, \tau, 0, t) = \alpha$ for $\sigma = 0$. $G(m, s, t)$ is the cumulative Gaussian distribution with mean m and standard deviation s at time t, given by

$$G(m, s, t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(t-m)^2}{2s^2}} dt.$$

Within this formalism, α represents the asymptotic value of the time-series. σ represents the gradient of the fitted function at the inflection point, this is where the derivative of the fitted function stops increasing and begins to decrease (for $G(m, s, t)$ this is the gradient of G at m). τ is the time taken to reach the stable regime, which is given by the intersection of α and the tangent of the fitted function at the inflection point. However, these fit parameters are not suitable for measurement of the adaptation period because they do not yield a value of relative change from fit start to end.

An example of fit is shown in Figure 3.4. The adaptation duration was defined as the difference between the time at the start of the fit, and the time at which the fitted curve was within 1% of the asymptotic value. The degree of adaptation was calculated as the difference between the proportion of synchronous responses at the start and end of the adaptation duration. If the fitted curve did not reach the 1% threshold within the trial, the adaptation was calculated as the difference between the time at the start of the fit and the end of the trial, and the degree of adaptation was calculated as the difference between the value of the fit at the beginning of the fit and at the end of the trial. Participants whose fit had an asymptotic value of less than 0.7 in the synchronous condition were removed from further analysis, as this may indicate a misunderstanding of the experimental instructions since for 0ms SOA trials, participants should respond synchronous throughout the entire trial.

3.2.2 Results

1 participant was removed from further analysis because the value of their asymptotic fit in the synchronous condition was less than 70% (representing 5% of the cohort). Participant's responses to the synchronous stimui are shown in Figure 3.5. As expected, the proportion of synchronous responses increases sharply at the beginning of the trial, and remains high throughout. The average time-course of participants responses for the remaining SOAs are shown in Figure 3.6. After the change to asynchrony, each SOA yields a different response curve. For some SOAs, adaptation occurs over a variable epoch as responses recover to perceived synchrony.

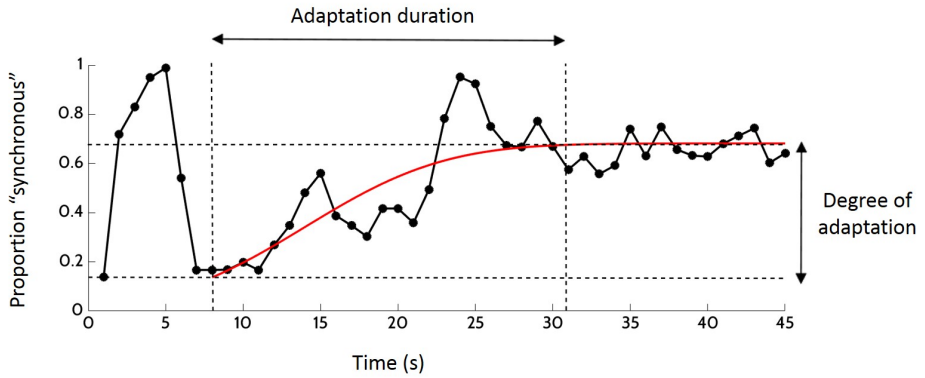

Figure 3.4: An example adaptation data set has been down-sampled to 45 points (black dots). The fit begins at 8s and tends asymptotically towards a value of around 0.7. The time at which the fitted curve reached a 1% threshold of the asymptotic value was used as the time at which adaptation ceased (rightmost vertical dashed line). The time difference between the start of the fit (leftmost vertical dashed line) and the end of the adaptation period was recorded as the adaptation duration. The degree of adaptation was calculated as the difference between the fitted values at the start and end of the adaptation duration (horizontal dashed lines).

In each figure, there is an initial steep rise in synchronous responses owing to the first 5s of physically synchronous stimuli. It appears that participants are less likely acheive 100% synchrony in the first 5s for asynchronous trials. To test this, I ran a one-way repeated-measures ANOVA on the percentage of synchronous responses after 5s by SOA. There was no significant effect of trial SOA on perceptual synchrony at 5s ($F(6, 108) = 1.844, p = 0.097, \eta_p^2 = 0.093$). This verifies participants were not cued to the trial asynchrony before the stimulus change. The mean proportion of synchronous responses after 5s was 82.69%, which may be indicative of general uncertainty about the synchrony of the stimuli or the experimental task.

The degree of adaptation that each participant demonstrated for each SOA was calculated. The average degree of adaptation for each SOA is depicted in Figure 3.7.

To investigate the effect of SOA on the degree of adaptation, a one-way ANOVA was performed on the degree of adaptation with SOA as a repeated-measure. There was a strong main effect of SOA ($F(6, 108) = 7.198, p < 0.0005, \eta_p^2 = 0.286$). Linear contrasts demonstrated an increase in the degree of adaptation as the trial SOA increased from auditory-first to visual-first ($F(1, 18) = 35.154, p < 0.0005, \eta_p^2 = 0.661$).

One-sample t-tests showed that the degree of adaptation was significantly greater than zero when the trial SOA was -100 ms, 200ms, 300ms, and

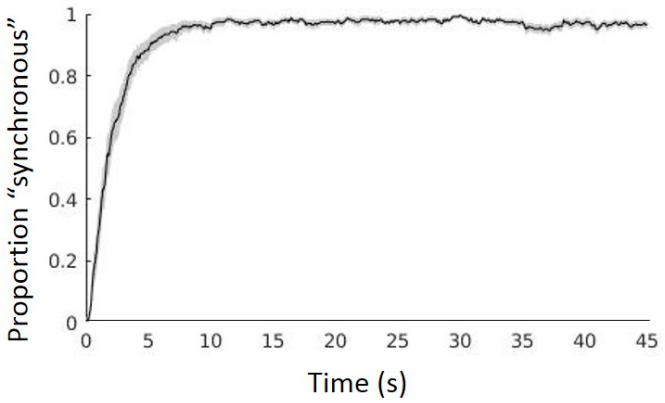

Figure 3.5: Proportion synchronous trial time-course with 0ms SOA for all participants ($n = 19$) in Experiment 1. The grey area represents the standard error of the means. The proportion of synchronous responses is consistently close to 1 after participants have started to respond. Note that the lowest value of the response curve between 5 and 20 is close to the value at the end of the trial, therefore the degree of adaptation at this SOA is low.

400ms ($t(18) = 3.14, p = 0.003; t(18) = 3.564, p = 0.001; t(18) = 12.583, p < 0.0005; t(18) = 8.771, p < 0.0005$ respectively).

I calculated the duration of adaptation for each participant as outlined in subsection 3.2.1. However, many of the fits obtained with this procedure did not reach their asymptotic value within the trial, and therefore these results were constrained by the fitting procedure and not fit for further analysis.

3.3 Experiment 2

The results of Experiment 1 suggested that the degree of adaptation is dependent on the size and direction of the trial asynchrony. In this experiment, I investigated the effect of extended exposure to physically synchronous stimuli on the time-course of adaptation. I predicted that the extended reference to physical synchrony would negatively impact the degree of adaptation.

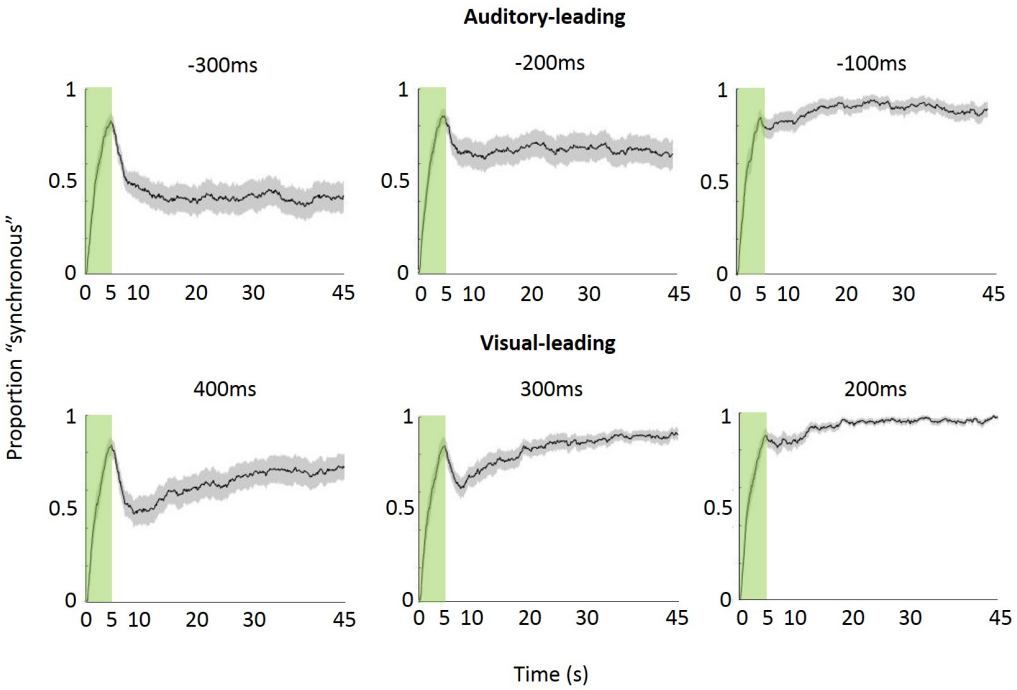

Figure 3.6: Mean proportion synchronous time-courses for all non-zero SOA conditions ($n = 19$). Grey areas represent the standard error of the means. The synchronous period at the beginning of each trial is shown in green. Each SOA yields a different response curve after the initial synchronous response. In each figure, there is a steep rise in the proportion of synchronous responses during the initial synchronous phase. At -200ms and -300ms, the proportion of synchronous responses continued to decrease after 5s. At -100ms there is a small dip in synchronous responses after 5s, after which synchronous responses increase. At 200ms, 300ms, and 400ms, the same pattern of a dip and subsequent rise is more pronounced.

3.3.1 Methods

Participants

17 naïve healthy adults took part in the experiment (mean age 19.9, range 3 years; 13 female). Participants responded voluntarily to advertisements and were given a small monetary compensation for their time. Participants gave written informed consent to the study, and to the use of their data. The study was approved by the Newcastle University Ethical Committee and was run in accordance with the Declaration of Helsinki.

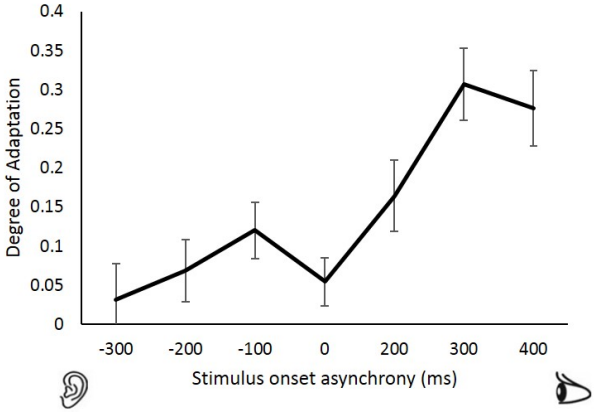

Figure 3.7: The degree of adaptation as a function of SOA for $n = 19$ participants in Experiment 1. Error bars represent the standard error of the means. The degree of adaptation is larger for visual-leading asynchronies than auditory-leading asynchronies. These values correspond to the time-series in Figure 3.6 and Figure 3.5 by the degree responses recovered after the initial change to asynchrony.

Stimuli

Stimuli for this experiment were the same as in Experiment 1 –Methods.

Procedure

Participants were presented with videos that began with 20s of physically synchronous auditory-visual stimuli, and then either remained synchronous, or changed to one of the following 6 SOAs: -300ms, -200ms, -100ms, 0ms, 100ms, 200ms, 300ms (as in Experiment 1). There were 12 trials for each asynchrony, leading to 84 trials of 50s each (*see* Figure 3.8).

Analysis

Because the auditory and visual components of the stimulus became asynchronous after 20s rather than 5s, points between 20s and 35s were considered to be suitable points to begin the final fit (*see* subsection 3.2.1). Otherwise analytical methodology is the same as in Experiment 1.

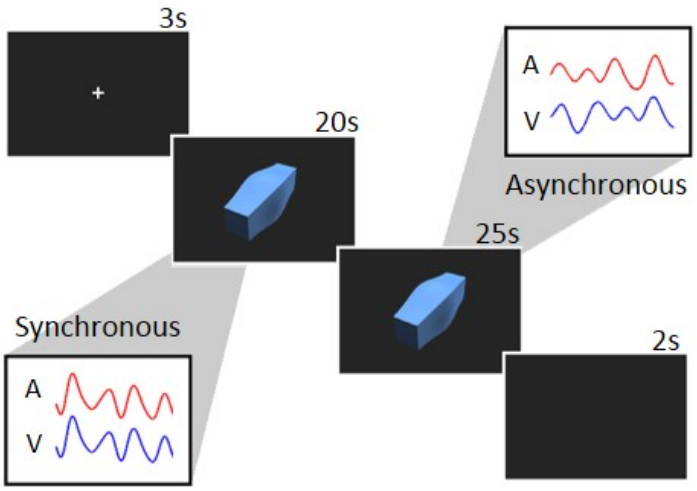

Figure 3.8: Experiment 2 trial time-course. Each trial began with with a fixation cross for 3s. The auditory-visual stimulus was then presented synchronously for 20s each trial. The synchrony of the stimulus then changed to the trial SOA for 30s. Participants continuously monitored the stimulus and reported their perceived simultaneity of the auditory and visual components. After stimulus presentation, the screen was blank for 2s before the onset of the next trial.

3.3.2 Results

Two participants were excluded from further analysis due to low asymptotic fit values in the synchronous condition (*see* subsection 3.2.1) representing 12% of the cohort. The average time-course of participants responses to synchronous stimuli are shown in Figure 3.9. Responses to the remaining SOA conditons are shown in Figure 3.10. After the change to asynchrony, each SOA yields a different response curve.

The average degree of adaptation for each SOA is presented in Figure 3.11. As in Experiment 1, there is an asymmetry between auditory-leading and visual-leading asynchronies. A one-way repeated measures ANOVA revealed a main effect of SOA ($F(2.64, 36.964) = 4.618, p = 0.01, \eta_p^2 = 0.248$). Linear contrasts demonstrated an increase in the degree of adaptation as the asynchrony changed from auditory-leading to visual-leading ($F(1, 14) = 7.583, p = 0.016, \eta_p^2 = 0.351$). One-sample t-tests showed that the degree of adaptation was significantly greater than zero when the SOA was 200ms ($t(14) = 2.282, p = 0.0193$) but not for other SOAs after correction for multiple comparisons.

As in section 3.2, I calculated the duration of adaptation for each participant. Again, these results were constrained by the fitting procedure and not fit for further analysis.

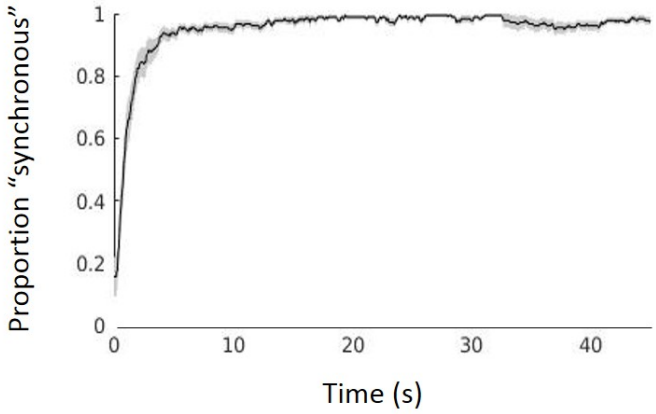

Figure 3.9: Proportion synchronous trial time-course with 0ms SOA for all participants ($n = 15$). The proportion of synchronous responses is consistently close to 1 after participants have started to respond. As expected, responses have a similar profile to the 0ms SOA condition in Experiment 1 (*see* Figure 3.5).

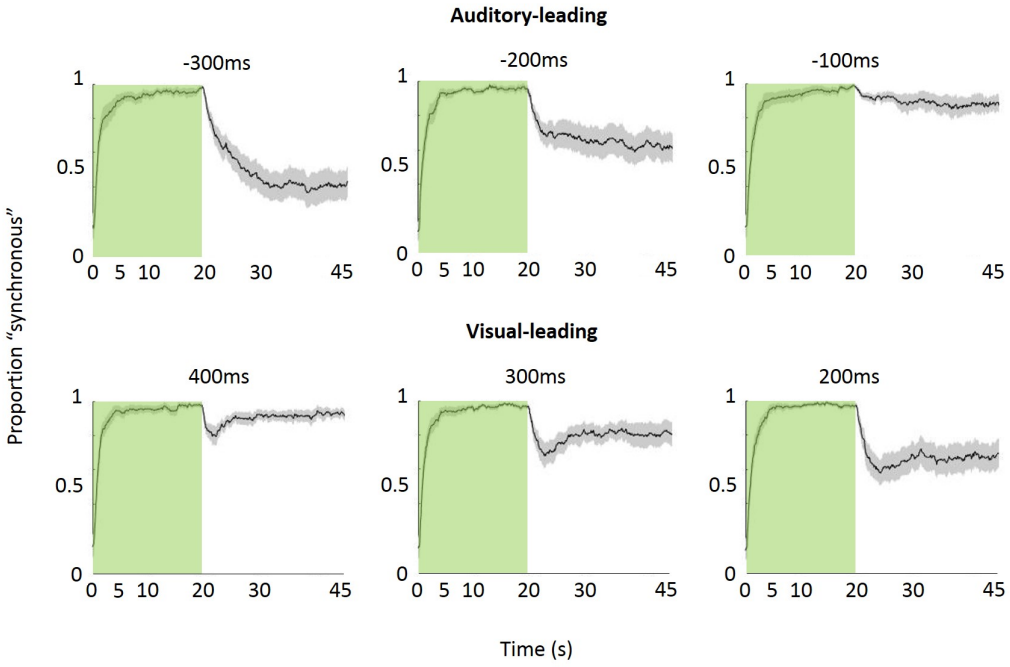

Figure 3.10: Proportion synchronous time-courses for all non-zero SOA conditions ($n = 15$). Grey regions represent the standard error of the mean value. Green regions represent the initial stage of the trial, during which the SOA was 0ms. Each SOA yields a different response curve after 20s. For visual-leading SOAs, the synchronous responses tend to increase over time and plateau, whereas responses to auditory-leading SOAs do not.

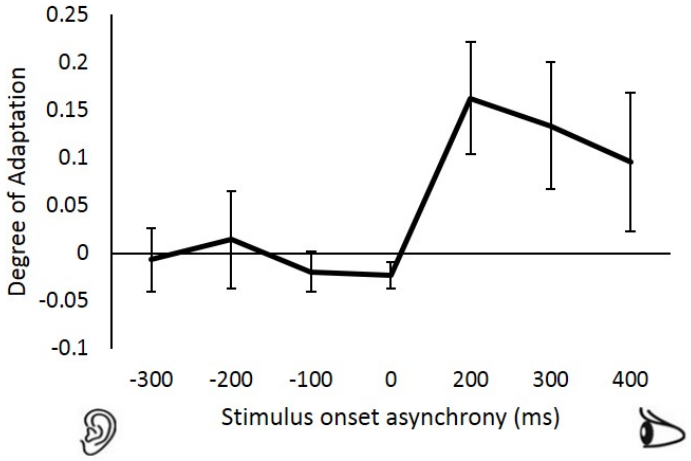

Figure 3.11: Degree of adaptation, the difference between the proportion of synchronous responses at the start and end of the fitted curve, as a function of SOA. Again Error bars represent the standard error of the mean values.

3.4 Experiment 3

In both Experiment 1 and Experiment 2, the degree of adaptation was asymmetrical, and was larger for visual-leading SOAs. The overall degree of adaptation for visual-leading SOAs was less in Experiment 2 (mean 0.15) than in Experiment 2 (mean 0.3). To further investigate how the duration of exposure to the synchronous stimuli affects subsequent adaptation, the above experiments were combined into a 2 $\times$ 2 (SOA $\times$ synchronous duration) factorial design.

3.4.1 Methods

Participants

20 naïve healthy adults took part in the experiment (mean age 20.8, range 4, 14 female). Participants responded voluntarily to advertisements and were given a small monetary compensation for their time. Participants gave written informed consent to the study, and to the use of their data. The study was approved by the Newcastle University Ethical Committee and was run in accordance with the Declaration of Helsinki.

Stimuli

See subsection 3.2.1.

Procedure

I compared two SOAs: 300ms, and -300ms, preceded by one of two synchronous durations: 5s and 20s (*see* Figure 3.12). Additionally, I kept the 0ms SOA condition as a control condition (*see* Experiment 1 Methods - *Analysis*). This led to a 2 × 2 factorial design (SOA × synchrony duration). I did not include the 0ms SOA in the ANOVA. The stimuli had a duration of 50s. There were 12 trials for each condition, leading to 60 trials of 55s each.

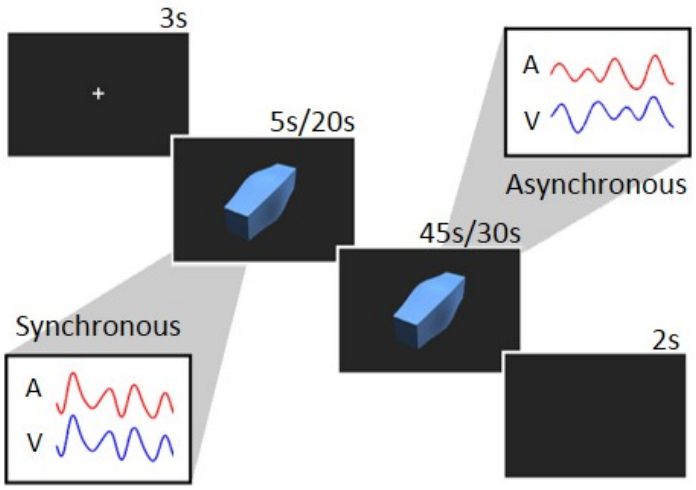

Figure 3.12: Experiment 3 trial time-course. Each trial began with with a fixation cross for 3s. The auditory-visual stimulus was then presented synchronously for either 5s or 20s each trial. The synchrony of the stimulus then changed to the trial SOA until the total duration of the stimulus was 50s. Participants continuously monitored the stimulus and reported their perceived simultaneity of the auditory and visual components. After stimulus presentation, the screen was blank for 2s before the onset of the next trial.

Analysis

See subsection 3.2.1 - *Analysis*.

3.4.2 Results

After fitting, 2 participants were removed from further analysis due to their response in the synchronous condition (10% of cohort). The average time-course of participants responses to synchronous stimuli are shown in Figure 3.13. Responses to the remaining SOA conditons are shown in Figure 3.14.

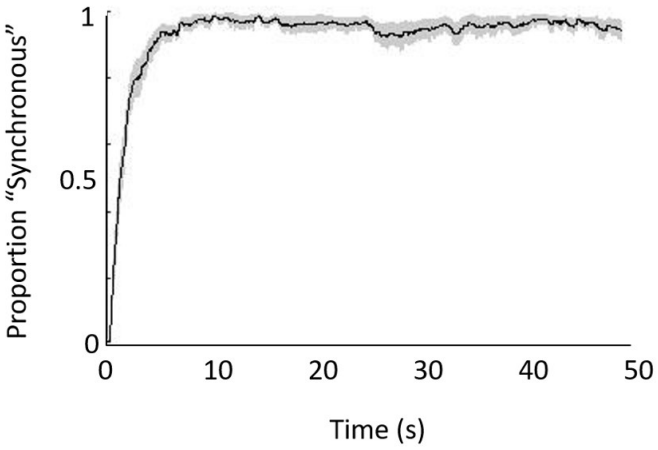

Figure 3.13: Proportion synchronous trial time-course with 0ms SOA for all participants ($n = 18$). The proportion of synchronous responses is consistently close to 1 after participants have started to respond. As expected, responses have a similar profile to the 0ms SOA condition in both Experiment 1 Experiment 2.

The degree of adaptation as a function of SOA and duration of the preceding synchronous period is shown in Figure 3.15. A 2 × 2 (SOA × synchronous duration) repeated measures ANOVA was performed on the degree of adaptation. There was a main effect of SOA ($F(1, 17) = 13.395, p = 0.002, \eta_p^2 = 0.441$) but no effect of the duration of the synchronous period ($F(1, 17) = 1.162, p = 0.296, \eta_p^2 = 0.064$). One-sample t-tests showed that the degree of adaptation was significantly greater than zero for visual-leading conditions (visual-leading 5s synchrony: $t(17) = 3.623, p < 0.0005$; visual-leading 20s synchrony: $t(17) = 2.707, p = 0.007$). Additionally, the degree of adaptation was significantly less than zero when the asynchrony was auditory leading and the adaptation duration was 20s (auditory-leading 5s synchrony: $t(17) = 1.103, p = 0.142$; auditory-leading 20s synchrony: $t(17) = 2.563, p = 0.010$; synchrony: $t(17) = 1.703, p = 0.52$).

Results pertaining to the adaptation duration have been omitted as they were heavily biased by the measurement procedure.

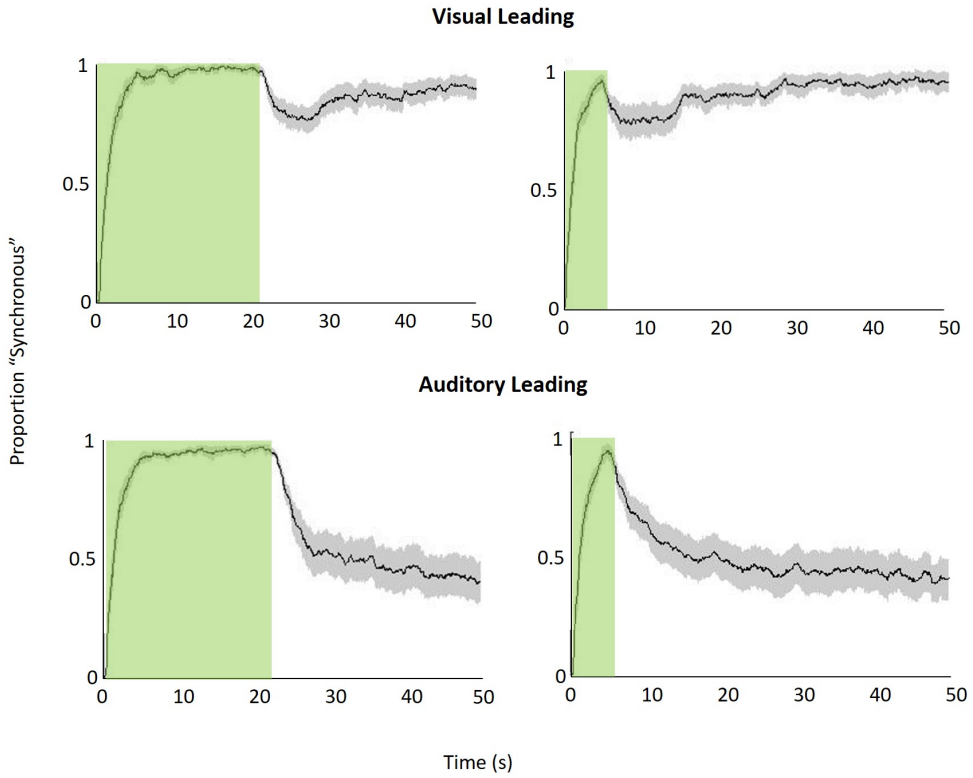

Figure 3.14: Proportion synchronous time-courses for all non-zero SOA conditions ($n = 18$). Grey regions represent the standard error of the mean value. Green regions represent the initial stage of the trial, during which the SOA was 0ms. After the change to asynchrony, the proportion of synchronous responses to visual-leading SOAs increases over time, whereas responses to auditory-leading SOAs do not.

3.5 Discussion

In this chapter, I conducted three experiments about auditory-visual adaptation, which is a key methodology in classical recalibration studies. In Experiment 1, I looked at time-courses of auditory-visual adaptation. To provide a base-line response, each trial began with 5s of synchronous presentation. I found that the SOA has a significant linear effect on the degree of adaptation, and that the degree of adaptation was significantly greater than zero when the SOA was -100ms, 200ms, 300ms, and 400ms. In Experiment 2, I replicated the result that the SOA significantly affects the degree of adaptation. Finally, in Experiment 3 I again replicated the effect of SOA on the degree of adaptation.

In all experiments, I was not able to measure the adaptation duration accurately because, in many cases, our fitting procedure indicated that it was

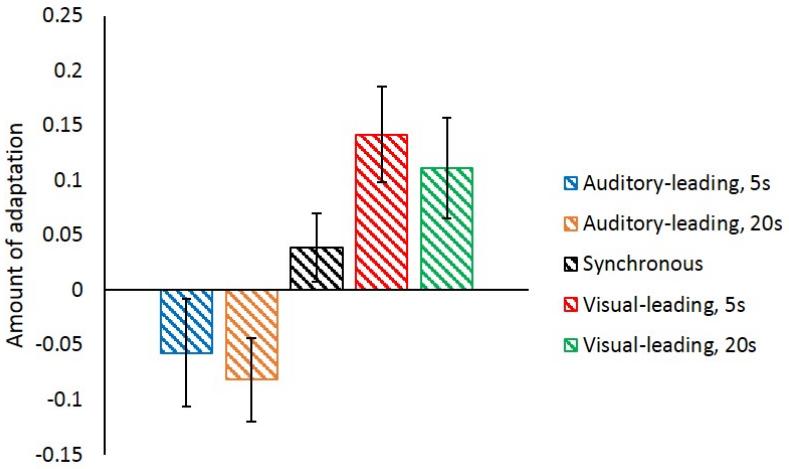

Figure 3.15: The degree of adaptation in Experiment 3 as a function of modality order and the duration of the initial synchronous presentation ($n = 18$). Adaptation occurred for visual-leading SOAs but not for auditory-leading SOAs.

longer than the trial length. While this may qualitatively indicate that adaptation can occurs over 30—40 seconds, I cannot verify this quantitatively. Future work on the time-course of adaptation should seek to refine the fitting procedure so that the adaptation duration can be measured accurately.

Classical recalibration studies have shown that prolonged exposure to asynchrony can shift the PSS in the direction of the exposed asynchrony [38]. These experiments had 3 minute extended adaptation periods before experimental blocks, as well as re-exposure periods before each trial [38, 40]. In this study, adaptation occurred within the 45s trial SOA and often over shorter 10-20 second epochs. Therefore, it is possible that PSS shifts presented in these classical recalibration experiments were driven primarily by the re-exposure phase. Additionally, this provides evidence that classical and rapid recalibration rely on different neural mechanisms, since adaptation cannot occur within a timescale applicable to rapid recalibration.

I hypothesised that adaptation would occur to a greater extent for visual-leading stimuli than for auditory-leading stimuli. This hypothesis was based upon the natural bias for visual-leading stimuli elicited by the multisensory system when judging synchrony or temporal order (*see* section 1.2 - Temporal Factors). I showed that although adaptation can occur for small auditory-leading asynchronies, the degree of adaptation was much greater for visual-leading asynchronies. For auditory-leading asynchronies larger than 100ms, adaptation did not occur. Interestingly, classical recalibration studies have demonstrated recalibration to auditory-leading stimuli with greater asynchronies, when adaptation may not have occurred. In these cases, it is

possible that recalibration effects in the auditory-leading direction were not driven by perceptual changes to the adaptation stimulus. This hypothesis may also describe the origin of the effect size modality order asymmetry in rapid recalibration.

There is a natural preference for visual leading asynchronies in nature, which may be driven by differences in physical and neural transduction speeds [27]. It is possible that low-level perceptual changes occur in the visual-leading direction only, whereas recalibration to auditory-leading asynchrony requires more time to establish higher-order associations. Overall, these experiments give context to the ongoing debate about differences between the underlying mechanisms of rapid and classical recalibration. Future work should address whether low-level neural correlates of adaptation (e.g. phase shifts of neural oscillations in primary sensory cortices [97]) occur more readily for visual-leading stimuli than auditory-leading stimuli.

Chapter 4

Neural Correlates of Adaptation to Temporal Asynchrony

In Chapter 3, I investigated temporal adaptation: the key methodological manipulation of sustained recalibration studies. I demonstrated an asymmetry in the adaptation effect, which challenges assumptions about how recalibration occurs.

To date, no brain regions have been associated with temporal recalibration or temporal adaptation. In this chapter, I will use the same stimuli as in Chapter 3 to manipulate the state of adaptation whilst collecting fMRI data. I will use continuous stimuli with some temporal characteristics similar to those of natural speech, to investigate the neural basis of adaptation to ecologically valid stimuli.

4.1 Introduction

Very little is known about how auditory-visual adaptation occurs in the brain. One recent study demonstrated a correlation between the size of the recalibration effect and the phase of neural oscillations in the auditory cortex [97]. The authors hypothesised that over time, ongoing cortical oscillations become entrained to external stimuli. Subsequently, precession occurs in the oscillatory phase to facilitate the encoding of the asynchronous event timing. However, it is not clear how or where this phase precession occurs, and

imaging with higher spatial resolution is necessary for further investigation.

Although little is known about temporal adaptation in the brain, many studies have addressed the neural responses to stimulus repetition. Typically, the neural response to some stimulus decreases with the number of repetitions [160–164] (for a review *see* [165]). This neural response is referred to as repetition suppression. One explanation for repetition suppression is that the brain is able to represent the presented stimulus using fewer resources when it is repeated. The sparse encoding of stimulus features leads to an increase in sensitivity to small changes that do not match the internal representation [161, 165, 166]. These models of stimulus repetition may apply to adaptation; in this case the neural response to temporal asynchrony will reduce as the duration of exposure increases.

For temporal adaptation, the concept of perceptual synchrony adds a confound not present in repetition suppression studies. This is because it is not clear whether existing patterns of neural activity are elicited via physical or perceptual synchrony or asynchrony. For example, in section 1.4: *Neural Correlates of Multisensory Integration* I discussed some brain regions which tend to be more active during temporal congruence than incongruence, and even display superadditive behaviours [14, 16, 105, 119]. However, these brain regions may not display these behaviours if both physical synchrony and asynchrony are perceptually synchronous to the observer, as a result of adaptation to prior stimuli.

One study has investigated the neural basis of perceptual synchrony and asynchrony [122]. In the scanner, participants were presented with asynchronous speech streams and asked to report when they felt that synchrony of the speech changed by pressing one of three keys: auditory-leading, synchronous, and visual-leading. These responses were then modelled as boxcar functions in the fMRI analysis. The authors demonstrated distinct regions within the superior temporal sulcus (STS) that were more active for perceptual synchrony than perceptual asynchrony *or vice versa*, and additionally identified a region in the prefrontal cortex (PFC) which was more active during perceptual asynchrony. These regions are thought to form a network in which the PFC modulates activity in the STS [122, 124, 126, 167]. It is possible that these brain regions will be involved in the underlying mechanism of adaptation, but it is not clear whether activity will reduce over time via repetition suppression.

To investigate, I used the novel paradigm and stimuli as described in Chapter 3 in conjunction with continuous fMRI. I manipulated the modality order of the asynchrony and the presentation order of synchronous and asyn-

chronous sections of the stimulus to elicit adaptation during specific periods. In addition, I measured perceptual synchrony throughout the trial time-course.

4.1.1 Aims

I hypothesised that:

- The overall neural response to the temporal asynchrony is increased when the asynchrony is first presented, relative to when the same asynchrony has been presented for a longer duration.
- Asynchronous perceptual states elicit a response in the superior temporal sulcus (STS) and prefrontal cortex (PFC).
- Synchronously presented stimuli will elicit a greater response in the STS than asynchronously presented stimuli.

4.2 Methods

4.2.1 Participants

23 right handed adults (14 female, 9 male; mean age 24, standard deviation 4.7, range 18–36) took part in the study. 23 participants reported normal or corrected to normal vision, 22 participants has normal hearing (one participant had moderate tinnitus but was able to hear the stimuli clearly). Participants were recruited via the Newcastle University Institute of Neuroscience volunteer mailing list and were given £5 per hour or their time. All participants gave informed consent to the experiment. The study was performed in accordance with the declaration of Helsinki and approved by the Ethics committee of Newcastle University.

4.2.2 Apparatus

The visual stimuli were back-projected onto a screen at the foot end of the scanner using a Canon XEED LCD projector (1280 $\times$ 1024 pixels, 60Hz).

Participants view the projection through an angled mirror attached to the head coil 10cm above their eyes. Auditory stimuli were presented with an MR-compatible audio system and delivered with in-ear electrostatic transducer headphones (NordicNeuroLab). To reduce scanner noise for participants, the headphones were modified such that they could be comfortably worn beneath ear defenders. Head motion was restricted by placing foam pads between the head and the head coil. The experiment was run on a Windows 7 PC. Psychophysics Toolbox version 3[1] [148, 149] was used to present stimuli and record responses (run on 32-bit MATLAB 2012, Mathworks, Inc.). Participants responded by key-press on a MR-compatible key pad (LumiTouch).

4.2.3 Stimuli

The auditory and visual stimuli used in this experiment are described in detail in Chapter 3. They consisted of a blue cuboid and an amplitude-modulated tone for which the modulation envelope was derived from speech. I was able to alter the delay between auditory and visual components of the stimulus smoothly, without any clipping, popping, or other indication. In this experiment, all auditory-leading asynchronies had an SOA of -300ms, and all visual-leading asynchronies had an SOA of 300ms. When auditory and visual components of the stimuli were synchronous, the SOA was 0ms.

I use the terms *physical synchrony* and *physical asynchrony* in this chapter simply as a means to differentiate from perceptual synchrony and asynchrony. Physical synchrony is used to describe the stimuli when the have been presented with an SOA of 0ms. I did not verify that the stimuli were in fact physically synchronous during the experiment, nor did I account for the time taken for signals to be transduced from the control PC to the projector and headphones. However, any deviations from the intended SOA were consistent across all trials and were orders of magnitude less than the experimental variation of SOA (300ms between condiditons).

4.2.4 Design

Each functional run consisted of 7 trials in which I presented the stimuli for 40s. In three trials, the synchrony of the stimuli was unchanged throughout presentation (auditory-leading, visual-leading, and synchronous). In the

[1] http://psychtoolbox.org/

remaining trials, I altered the physical synchrony of the presented stimuli after 20s, but always included 20s of synchrony (auditory-leading to synchronous, synchronous to auditory-leading, visual-leading to synchronous, and synchronous to visual-leading –*see* Figure 4.1). Participants continuously reported on the synchrony of the stimuli throughout each trial. Fixation crosses were presented for 20s before each trial, and after the last trial. The trial order was random for each run and there were were 4 functional runs per participant.

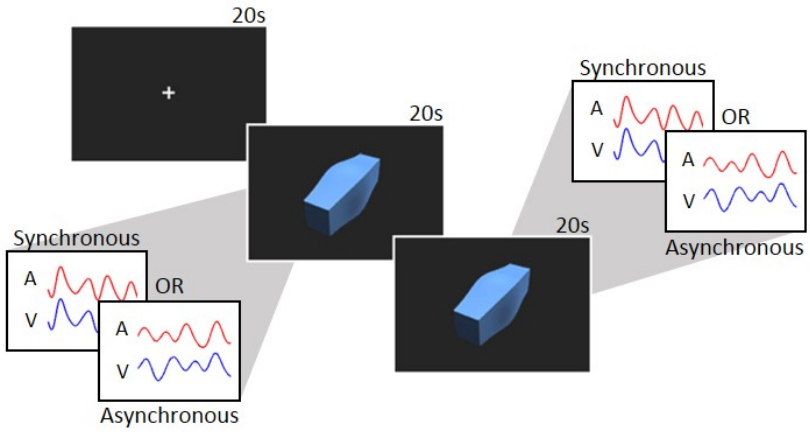

Figure 4.1: fMRI behavioural task. Each trial began with with a fixation cross for 20s. The auditory-visual stimulus was then presented for 40s each trial. In four out of seven conditions, the asynchrony of the stimulus changed after 20s for the remaining 20s of the trial. If the first interval was synchronous, the second interval was auditory-leading, visual-leading, or synchronous. If the first interval was visual-leading, the second interval was visual-leading or synchronous. Finally, if the first interval was auditory-leading, the second interval was auditory-leading or synchronous. Participants continuously monitored the stimulus and reported their perceived simultaneity of the auditory and visual components. The next trial began immediately after the offset of the stimuli.

4.2.5 Procedure

Index and middle fingers were used to report the perceived synchrony or asynchrony of the stimuli. Participants were asked to hold down a response key at all times whilst stimuli were presented, and to remove pressure during the fixation period. Participants were instructed to hold one response key while they perceived the stimuli to be synchronous, and to hold the other response key perceived the stimuli to be asynchronous. Outside the scanner, participants were familiarised with viewing and responding to the stimuli in a practise run (three 30s trials: auditory-leading, visual-leading, and synchronous). Response keys were counterbalanced between participants and

responses were captured every 17ms throughout each functional run. Additionally, participants were instructed not to respond during the fixation period.

Once inside the scanner, soft pads were placed in the head coil to restrict movement. I then presented the stimuli whilst performing dummy scans to ensure participants were able to comfortably hear the stimuli over scanner noise. Each functional run was 7 minutes 20s in duration. There were four functional runs for 22 participants and 3 for one participant due to technical issues.

4.2.6 Image Acquisition

All participants were scanned at the Newcastle Magnetic Resonance Centre. Anatomical T1-weighted images and functional T2*-weighted Echo Planar Images (EPIs) were acquired from a 3 T Phillips Intera Achieva magnetic resonance scanner using a Phillips 32-channel receive only head coil. The high-resolution T1-weighted scan consisted of 150 slices and took approximately 5 minutes to acquire. The parameters of the structural scan were: repetition time (TR) = 9.6ms, echo time (TE) = 4.6ms, flip angle = 8°. The field of view (FOV) was 240mm × 240mm × 180mm with a matrix size of 208 × 208 pixels. Each voxel was 0.94mm × 0.94mm × 1.2mm in size. The T2*-weighted Echo Planar Images (EPIs) were: acquisition time (TA) = 1.3s, TR = 2s, TE = 30ms. The FOV was 192mm × 192mm × 115mm in size with a matrix size of 64 × 62 pixels. Each voxel was 3mm × 3mm × 4mm in size, with a 1 mm gap between slices. I used sensitivity encoding (SENSE) with factor = 2 to increase the signal-to-noise ratio of the functional images. For each participant, a total of 230 functional images were acquired in each run (duration 7 minutes 20s). Before each functional run, 10 "dummy" scans were acquired to allow for calibration of the T1 signal.

4.2.7 fMRI Preprocessing

To correct for head motion, functional images were realigned and re-sliced to match the first image for each participant. These images were then normalised to a standard Montreal Neurological Institute (MNI) EPI T2*-weighted template with a re-sampled voxel size of 3 mm × 3 mm × 3 mm. Resulting images were spatially smoothed with a Gaussian kernel (6 mm full-width-at-half-maximum). To remove low frequency signal drifts, a high-pass

filter with a cutoff of 128 s was applied.

4.2.8 fMRI Whole-brain Analysis

For all analyses I initially preprocessed data using SPM8[2] [168]. I used the general linear model (GLM) with a two-step mixed-effects approach. A fixed-effects model was used to analyse each participant's data set. A random-effects model was used at the group level. There was no additional smoothing of the images at the group level. I analysed data at the first level using 3 separate design matrices, each to test a different hypothesis.

Adaptation

I constructed the first design matrix to address the first hypothesis: that overall neural responses to temporal asynchrony reduce over time. Recall that there were seven conditions: synchronous, auditory-leading, visual-leading, auditory-leading to synchronous, visual-leading to synchronous, synchronous to auditory-leading, and synchronous to visual-leading. Each condition was presented for 40s. To test the hypothesis, I divided the 40s duration into the first and second 20s interval. I modelled the onset of the second interval for each of the seven conditions as a boxcar function with a duration of 20s. I created three additional regressors for the onset of the first interval. The first regressor modelled the onset of the auditory-leading and auditory-leading to synchronous conditions (2 trials). The second modelled the onset of the visual-leading and visual-leading to synchronous conditions (2 trials). The third regressor modelled the onset of the remaining conditions (3 trials). Lastly, fixation periods were modelled as a single boxcar function. All boxcar functions were convolved with the canonical heamodynamical response function and entered into the design matrix as regressors of interest. In addition, the movement parameters and a constant term for each functional run were included in the design matrix as regressors of no interest. A linear combination of the regressors was fitted to the BOLD signal to estimate the beta weight for each regressor.

For the first-level analysis, contrast images were computed from the beta-weight images. I focused on the following contrasts of the second 20s intervals(but report other significant contrasts for completeness):

[2]https://www.fil.ion.ucl.ac.uk/spm/

- Visual adaptation: visual-leading vs. synchronous to visual-leading.
- Auditory adaptation: auditory-leading vs. synchronous to auditory-leading.

For the second-level group analysis, one-sample t-tests of participants' contrast images were conducted at each voxel. For statistical analyses, I used an initial threshold of $p_{unc} = 0.001$ and extent threshold of $k = 10$. I considered clusters as significant if $p < 0.05$ FWE-corrected at the cluster level (unless otherwise stated). I only report significant contrasts.

Perceptual Synchrony and Asynchrony

The second design matrix tested the second hypothesis: that asynchronous perceptual states elicit responses in the STS and PFC. Participants' responses were used to model periods of perceptual synchrony, perceptual asynchrony, and neither percept (i.e. no response). The onset and duration of these 3 perceptual conditions were based on participants' key presses, convolved with the canonical hemodynamic responses, and entered into the design matrix as regressors of interest. As before, the design matrix also included the movement parameters and a constant term for each functional run as regressors of no interest. A linear combination of the regressors was fitted to the BOLD signal to estimate the beta weight for each regressor. For the first level analysis, I contrasted perceptual synchrony against perceptual asynchrony. For the second-level group analysis, one sample t-tests of participants' contrast images were conducted at the voxel level.

Physical Synchrony and Asynchrony

Here I use the term "physically synchronous" to mean that the stimuli were presented with an SOA of 0ms, and "physically asynchronous" to mean that stimuli were presented with any non-zero SOA. I did not verify that the stimuli were absolutely physically synchronous using any measurement apparatus. This terminology serves only to differentiate this trial design, as it is possible for stimuli to be presented asynchronously but perceived synchronously. The final design matrix tested the hypothesis that physically synchronous stimuli elicit responses in the STS. For all 7 conditions, I modelled the entire 40s duration as 40s boxcar functions. Fixation periods were modelled as a single boxcar function. As before, the design matrix also included the movement

parameters and a constant term for each functional run as regressors of no interest. A linear combination of the regressors was fitted to the BOLD signal to estimate the beta weight for each regressor. For the first-level analysis, I focused on the following contrasts:

- Synchronous vs. (auditory-leading + visual-leading).
- Synchronous vs. auditory-leading.
- Synchronous vs. visual-leading.

For the second-level group analysis, one-sample t-tests of participants' contrast images were conducted at each voxel.

4.3 Results

4.3.1 Behaviour

One participant did not respond to the stimuli in any functional run. Another participant failed to respond during one functional run. I removed these runs from further fMRI analysis since I cannot be certain that these participants were attending to the stimuli during this time. The time-courses of participants responses from the remaining runs are shown in Figure 4.2 & Figure 4.3.

Due to the low number of trials for each condition, I was not able to adequately fit the behavioural data to verify adaptation effects at the participant level. However, note the progressive increase in the proportion of synchronous responses over time (*see* visual-leading stimuli in Figure 4.3) This is indicative of adaptation.

4.3.2 fMRI Whole-Brain Analysis

I used a liberal initial threshold of $p_{unc} = 0.001$ and $k = 10$ for exploratory analysis of the group level contrasts. Locations of significant clusters were mapped via the WFU Pickatlas Toolbox [169] using peak voxel MNI coordinates.

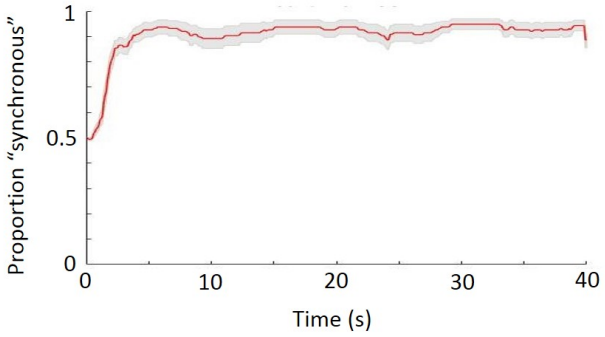

Figure 4.2: Time-courses of participant's responses to the multisensory stimuli during synchronous trials ($n = 22$). A synchronous response is assigned a value of 1, an asynchronous response the value of 0, and no response is assigned 0.5. As expected, the proportion of synchronous responses remains high throughout the trial (as in Chapter 3).

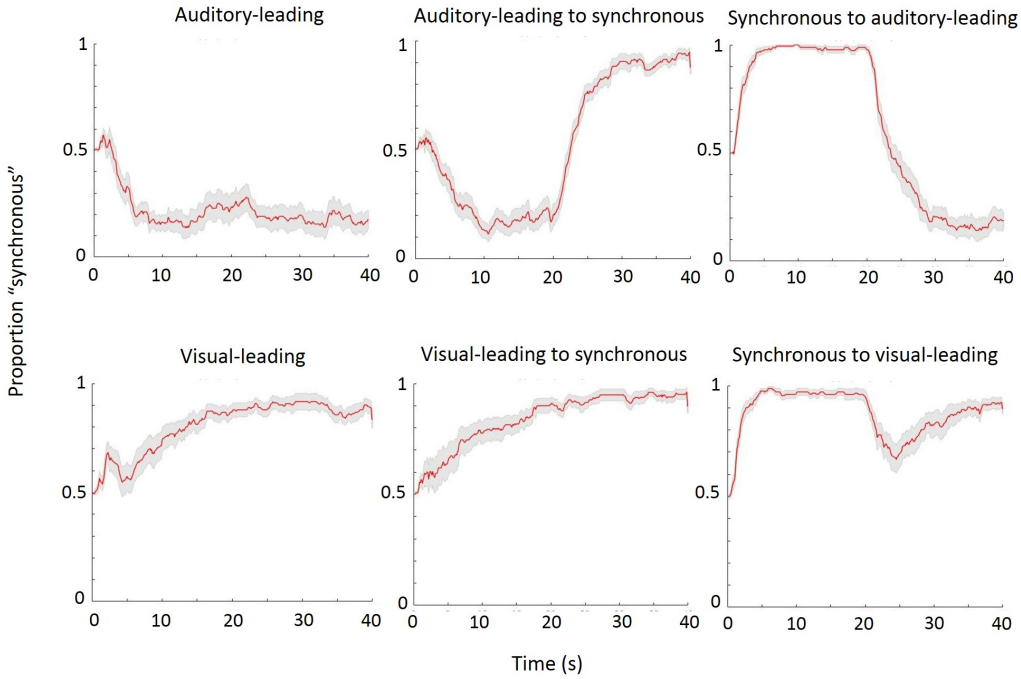

Figure 4.3: Time-courses of participant's responses to the multisensory stimuli during asynchronous trials ($n = 22$). A synchronous response is assigned a value of 1, an asynchronous response the value of 0, and no response is assigned 0.5.

Adaptation

I compared activation during the final 20s of trials in which the asynchrony was the same, but the asynchrony of the initial 20s was different. During visual-leading asynchrony, regions in the posterior cingulate cortex (PCC) and middle occipital gyrus (MOG) were significantly more active when the

preceding interval was visual-leading, compared to when the preceding interval was synchronous[3] (*see* Table 4.1, and Figure 4.4).

Location	Hem	k	Z	MNI coordinate			p_{unc}	p_{corr}
				x	y	z		
posterior cingulate cortex	L	66	4.37	-15	-61	4	$< 10^{-5}$	0.005
middle occipital gyrus	L	74	4.25	-9	-97	10	$< 10^{-5}$	0.003
posterior cingulate cortex	R	44	3.72	18	-61	10	0.003	0.029
k, cluster size; Z, Z-score; p_{unc}, uncorrected cluster-level p-value; p_{corr}, FWE-corrected cluster-level p-value								

Table 4.1: Visual-leading > synchronous to visual-leading ($n = 22$). Significantly greater activation occurred during visual-leading asynchrony when preceded by visual-leading than when preceded with synchrony. Coordinates are the local maxima for $p < 0.05$.

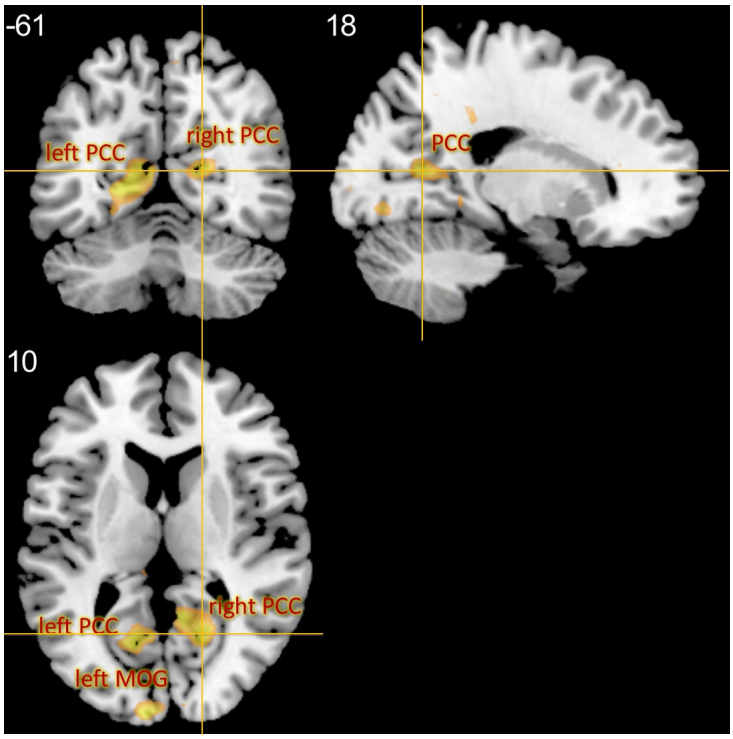

Figure 4.4: Visual-leading > synchronous to visual-leading ($n = 22$; final 20s). In bilateral PPC and left MOG, significantly greater activation occurred during visual-leading asynchrony when preceded by visual-leading than when preceded by synchrony.

There was significantly greater activation in the intraparietal sulcus (IPS) and middle temporal gyrus (MTG) when synchronous stimuli were preceded by auditory-leading stimuli compared to when preceded by synchronous stimuli (*see* Table 4.2 and Figure 4.5). There were no other significant activations.

[3]The left and right PCC clusters have their peak voxels and 71.4% of their volume inside the PCC when using the Talairach Daemon atlas.

Location	Hem	k	Z	MNI coordinate			p_{unc}	p_{corr}
				x	y	z		
intraparietal sulcus	R	311	4.55	57	-43	46	10^{-5}	10^{-5}
intraparietal sulcus	L	71	4.28	-45	-40	40	10^{-5}	0.004
middle temporal gyrus	R	55	4.23	63	-46	-11	0.001	0.015
k, cluster size; Z, Z-score; p_{unc}, uncorrected cluster-level p-value; p_{corr}, FWE-corrected cluster-level p-value								

Table 4.2: Auditory-leading to synchronous > synchronous ($N = 22$). Significantly greater activation occurred during synchrony when it was preceded by auditory-leading stimuli than when preceded with synchrony. Coordinates are the local maxima.

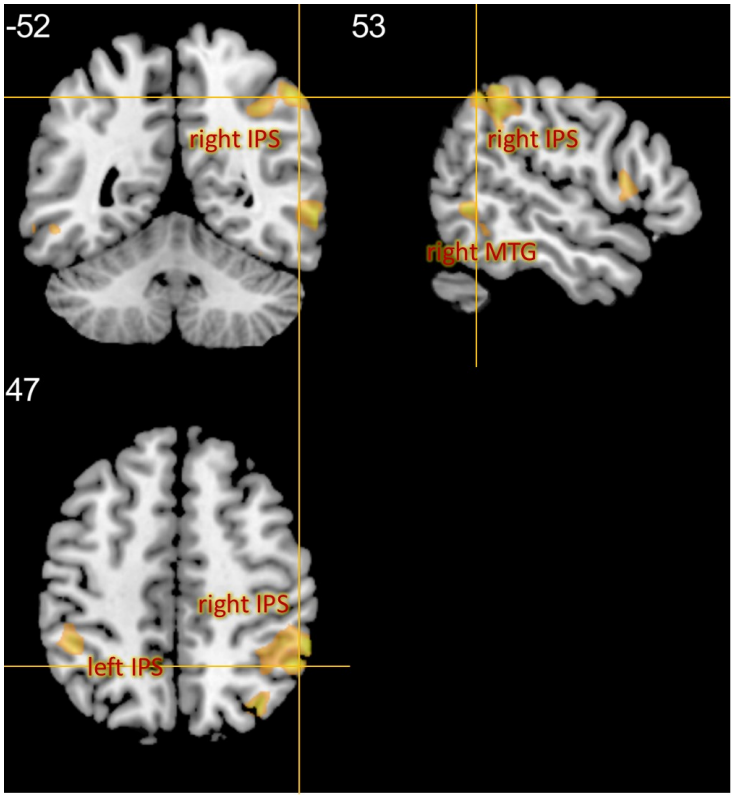

Figure 4.5: Auditory-leading to synchronous > synchronous ($N = 22$). In bilateral IPS and right MTG, significantly greater activation occurred during synchrony when it was preceded by auditory-leading stimuli than when preceded with synchrony.

Perceptual Synchrony and asynchrony

I contrasted perceptual synchrony against perceptual asynchrony as reported by participants during each trial, both correct and incorrect responses are included (*see* subsection 4.2.4). The average perceptual timecourses are depicted in Figure 4.2 and Figure 4.3. Significant bilateral activation occurred in frontal and parietal areas during perceptual asynchrony compared with perceptual synchrony (*see* Table 4.3 and Figure 4.6).

Location	Hem	k	Z	MNI coordinate			p_{unc}	p_{corr}
				x	y	z		
inferior frontal gyrus	R	418	5.42	33	23	-8	$< 10^{-5}$	$< 10^{-5}$
inferior frontal gyrus	L	140	4.71	-33	20	-8	$< 10^{-5}$	$< 10^{-5}$
intraparietal sulcus	R	240	4.50	60	-43	43	$< 10^{-5}$	$< 10^{-5}$
intraparietal sulcus	L	49	4.20	-42	-52	52	0.003	0.032
inferior frontal gyrus	R	47	3.82	48	47	1	0.004	0.037
k, cluster size; Z, Z-score; p_{unc}, uncorrected cluster-level p-value; p_{corr}, FWE-corrected cluster-level p-value								

Table 4.3: Local maxima for perceptual asynchrony > perceptual synchrony ($N = 22$) (*see* Figure 4.6)

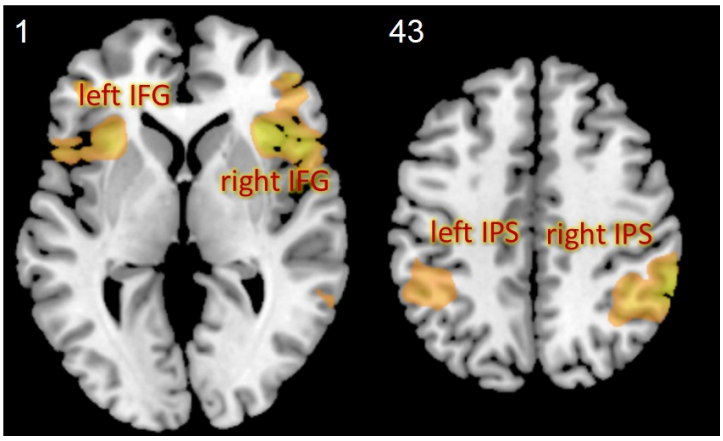

Figure 4.6: Perceptual asynchrony > perceptual synchrony ($N = 22$). In bilateral IPS and IFG, significantly greater activation occurred during perceptual asynchrony than during perceptual synchrony.

Physical synchrony and asynchrony

I investigated contrasts between the synchronous, auditory-leading, and visual-leading conditions. Three regions were significantly more active during auditory-leading stimuli than during synchrony (*see* Table 4.4, Figure 4.7), but no regions were more active for synchrony than asynchrony. No significant clusters were present for any visual-leading contrast.

Location	Hem	k	Z	MNI coordinate			p_{unc}	p_{corr}
				x	y	z		
inferior frontal gyrus	R	277	4.40	39	26	-11	$< 10^{-5}$	$< 10^{-5}$
inferior frontal gyrus	L	86	4.11	-36	20	-11	$< 10^{-5}$	0.001
intraparietal sulcus	R	60	3.79	48	-46	40	0.001	0.009
k, cluster size; Z, Z-score; p_{unc}, uncorrected cluster-level p-value; p_{corr}, FWE-corrected cluster-level p-value								

Table 4.4: Auditory-leading > Synchronous stimuli ($N = 22$) (local maxima for $p < 0.05$, *see* Figure 4.7).

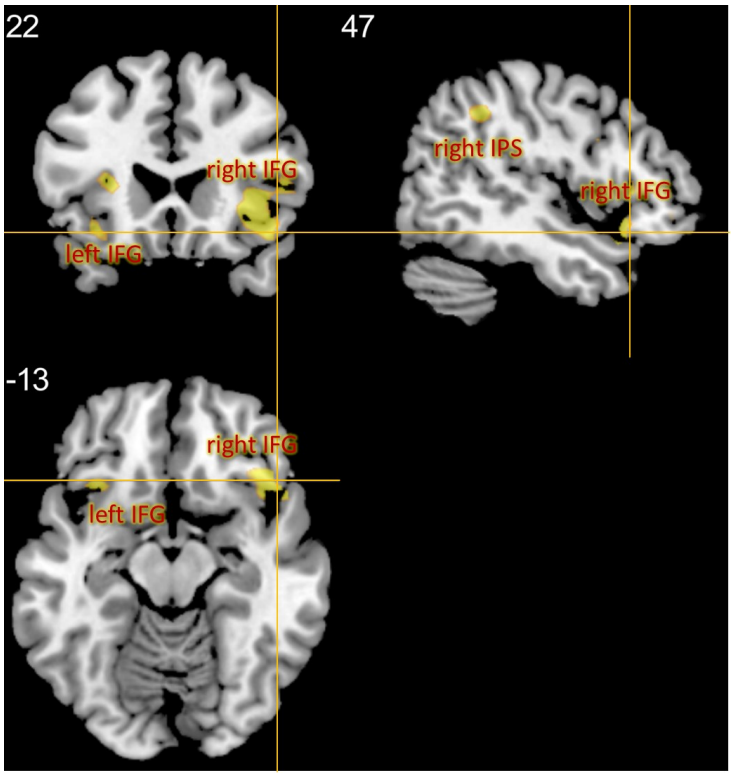

Figure 4.7: Auditory-leading > Synchrony ($N = 22$). There was significantly greater activation in bilateral IFG and right IPS during auditory-leading stimuli than during synchronous stimuli.

4.4 Discussion

The neural basis of auditory-visual temporal adaptation has not yet been elucidated. I developed a novel stimulus set which has allowed me to study temporal adaptation in the brain, as well as gain insight to the basis of physical and perceptual asynchrony. My analysis of brain activation during physical and perceptual synchrony and asynchrony has yielded significant bilateral clusters in frontal and parietal regions associated with multisensory processing. Analyses concerning adaptation yielded bilateral clusters in the posterior cingulate cortex (PCC). Counter to my hypothesis, which was based on repetition suppression, these clusters were more active when participants had adapted to the asynchronous prior rather than during adaptation, but only for visual-leading stimuli.

Throughout each trial, participants were asked to monitor the stimuli, and report on the synchrony throughout presentation. Figure 4.3 depicts the average time-course of the participants for each condition. I did not perform per participants analyses on these results due to the low number of trials. How-

ever, the characteristics of group level response time-courses are as expected, showing similar characteristics to the responses reported in Chapter 3. When the visual-leading asynchrony was preceded by synchrony, there is an initial dip in the synchronous response followed by a recovery phase in which participants adapt to the new asynchrony. As in Chapter 3, adaptation appeared to occur for visual leading stimuli only, and not for auditory leading stimuli.

I found that activation in the inferior frontal gyrus (IFG) and intraparietal sulcus (IPS) was increased during physically auditory-leading stimuli compared with physically synchronous stimuli, and during perceptual asynchrony compared with perceptual synchrony. These regions have been reported in studies of auditory-visual integration using various stimuli and tasks [106, 107, 121, 122, 124, 126, 170, 171]. However, with the exception of Noesselt et al. [122], these studies reported increased activation during synchrony, not asynchrony. It is possible that this is due to stimulus manipulations carried out by Noesselt et al. as well as the present study, which increased stimulus ambiguity. For example, Noesselt et al. [122] manipulated the synchrony of the stimuli presented to each participant to induce an equal number of perceptual switches between synchrony and asynchrony. Calvert at al. [16] also presented participants with auditory-visual speech, but the visual component of the stimulus was dubbed with a completely uncorrelated, incongruent auditory-component. Perhaps the increased ambiguity present when auditory-visual streams are asynchronous but associated in time, rather than completely incongruent, increases cognitive load due to perceptual decision making. In [171], activation in the frontal and parietal regions were observed whilst participants performed temporal order and simultaneity judgements. In this case, activation of the IPS and IFG was more robust when I contrasted perceptual states than when I compared physical asynchronies, suggesting that these regions are more associated with judgements about synchrony than synchrony itself. More research is needed to dissociate these effects.

With regard to activation in bilateral IFG during auditory-visual asynchrony, it has been proposed that perceptual asynchrony requires concurrent manipulation of separate memory representations in the prefrontal cortex [122]. Assessments of the reliability of each stimulus modality may be performed in frontal areas in order to provide topt-down modulation of sensory processing [124]. In my experiment, participants were asked to monitor the intentionally ambiguous synchrony of the stimuli, but it is not clear whether increased BOLD responses in frontal regions are associated with stimulus ambiguity, task demand or both. It would be beneficial to repeat this experiment with a stimulus irrelevant task to determine the extent to which this frontal activation is reliant on task demands. For example, participants could be asked to passively view the presented stimuli whilst simltaneously

completing a target detection task (i.e. responding to the detection of a target, say a red dot presented above the cuboid). This may make clear whether frontal activation is related to conscious decisions about synchrony, as was required by the current paradigm.

I hypothesised that the BOLD signal should increase when the synchrony of the stimulus is changed, as the asynchrony represents a new novel stimulus feature. I hypothesised that the BOLD response to asynchrony should not be large when the asynchrony has been presented over an extended duration. When I presented visual-leading asynchronous stimuli, there was significantly more activation in the PCC when the stimulus was preceded by the same visual-leading asynchrony, rather than by a synchronous stimulus. This increased activity may be representative of the brain response when it is adapted to, or compensating for, a stimulus with a visual-leading asynchrony. This result is contradictory to stimulus priming studies, in which repeatedly presented stimuli elicit response suppression [160–164]. It is likely that this results from the underlying perceptual uncertainty of the temporally asynchronous stimulus

Very little is known about the PCC, which forms a key node of the Default Mode Network, wherein each node displays larger BOLD responses during rest than during cognitive tasks. The PCC and ACC are highly anatomically [172] and functionally [173] connected brain regions. In addition, the ACC is known to respond to conflicts in information processing [174], and in turn redirects attention to the most behaviourally relevant stimuli to assist with cognitive control [175]. Evidence is emerging that the PCC "switches mode": functional connectivity changes dynamically in response to task demands, suggesting it has a roles in cognitive control and conflict monitoring as well as the default mode network [176]. Finally, recent work with non-human primates suggests that the PCC monitors environmental change to signal behavioural change [177]. In the context of my experiment, the PCC is an ideally placed hub [178] for functional connectivity between the intraparietal sulcus and inferior frontal gyrus [173], and may have been involved in monitoring the ongoing temporal conflict to enact behavioural change. To further investigate the role of the PCC in temporal adaptation, a future study could examine whether functional connectivity between the PCC and IPS or IFG can be modulated by the state of adaptation.

When participants were presented with synchronous stimuli, there was significantly more activation in parietal and temporal areas when this was preceded by an auditory-leading, asynchronous stimulus rather than a synchronous stimulus. This may be indicative of temporal recalibration. However it should be noted that these significant regions of activity were also present

when contrasting physically auditory-leading stimuli against synchrony. The behavioural response that occurred after changing from an auditory-leading stimulus to a synchronous stimulus took around 10s to stabilise (*see* Figure 4.3. Therefore, it is also possible that BOLD responses remained elevated due to the previous asynchrony. To more closely investigate adaptation aftereffects, a future study should compare the decay of the BOLD response in the IPS against the average BOLD response in sensory regions.

In summary, I have applied the robust behavioural effects present in Chapter 3 to manipulate behavioural states and measure the resulting neural response. This work supports the hypothesis that frontal and parietal regions contribute to the perception of asynchrony and are likely functionally linked in this context. In light of these findings, it is important to note that the presented stimuli were derived from speech, but held no semantic information. This work also indicates that the PCC may act as a monitor and facilitator of temporal adaptation, and more work is needed to discover how this is accomplished.

Chapter 5

Temporal Congruency and Correlates of Resting GABA Concentration

The inhibitory neurotransmitter gamma-aminobutyric acid (GABA) may have multifaceted effects on sensory integration in the brain. In this chapter, the strength of auditory-visual integration is measured using a robust behavioural paradigm [179], and measure resting GABA in two brain regions using magnetic resonance spectroscopy (MRS). I aim to uncover the extent to which inhibitory neurochemical mechanisms contribute to integration.

5.1 Introduction

Inhibition is an important factor in sensory integration. Many studies have shown that inhibition, mediated by the neurotransmitter gamma-aminobutyric acid (GABA), is important for sensory performance. For example, Wolf et al. [128] measured the spatial selectivity of neurons in the primary visual cortex of a kitten, before and after administering a GABA antagonist. The spatial selectivity of the neurons decreased after administration of the antagonist, demonstrating the role of GABAergic inhibition in sensory tuning.

Recently, methodology has been developed to measure the concentration of GABA *in vivo*. The concentration of the inhibitory neurotransmitter gamma-aminobutyric acid (GABA) is correlated with visual orientation discrimination performance [142]. The authors suggested that perception is improved

by the presence of GABAergic inhibition, which leads to more selectively tuned neural responses. GABA may inhibit neural responses in nearby receptive fields to improve spatial tuning [180].

Another recent investigation demonstrated that GABA concentration is important for tactile discrimination. In the experiment, participants were presented with standard and comparison vibrotactile stimuli in sequence, and judged whether the frequency of the second stimulus was the same or different to the the first. An adaptive procedure was used to estimate the 75% correct threshold. There was correlation between GABA concentration in the sensorimotor cortex, and individual differences in tactile discrimination performance [143]. The authors postulated that inhibition is generally important in sensory discrimination performance.

A general function of multisensory integration is to suppress neural responses to stimuli that are not aligned in space and time [8, 15, 16, 91, 98, 100, 101, 127] (for a review *see* [2]). Given the importance of inhibition for neuronal selectivity, it is likely that inhibition plays a role in multisensory integration.

Regarding spatial selectivity, the role of inhibition in multisensory integration is less clear because different modalities represent space in different coordinate spaces. Some transformation must occur to align the reference frames of each modality, to perform integrative operations. There is evidence that this transformation occurs in the parietal cortex [115, 116] and superior colliculus (SC) [181]. Computational models have suggested that inhibitory signals are important for mainting spatial frames of reference between modalities [182, 183]. It is possible that GABAergic inhibition mediates the suppression of neural responses to multisensory stimuli that are not aligned in space.

Neural oscillations likely form the basis upon which we encode temporal structure [59, 61, 97, 184]. Fast spiking GABAergic interneurons are important to the generation of high-frequency gamma oscillations [185, 186]. Both [142] and [143] demonstrated a ternary correlation between the concentration of GABA, behavioural performance, and gamma band power. This suggests that GABAergic inhibition may play a role in temporal discrimination.

A recent experiment has demonstrated a ternary correlation between GABA, gamma band power, and the illusion rate of the sound induced flash illusion [146]. This experiment demonstrates that GABA can affect multisensory percepts. However, this illusory percept is not indicative of how multisensory integration enhances sensory perception in general. I aimed to demonstrate that inhibition mediated by GABA plays a general role in multisensory enhancement. To do so, I utilised an auditory discrimination paradigm which

elicits a robust effect [179], and compared behavioural results with GABA concentration in two multisensory regions. I identified these multisensory regions by their increased BOLD response to temporal congruence in a previous fMRI paradigm using the same stimuli [126].

5.1.1 Aims

I hypothesised that:

- The extent to which integration augments performance in an auditory discrimination task is correlated with resting GABA concentration in the superior temporal sulcus (STS) and intraparietal sulcus (IPS).

5.2 Methods

5.2.1 Participants

16 healthy adults (7 male, mean age 21.8, range 18–39, 1 left handed) took part in the experiment. All participants reported normal or corrected to normal vision and hearing and were naïve to the experiment and stimuli. Participants responded voluntarily to advertisements and were given a small monetary compensation for their time. Participants gave written informed consent to the study, and to the use of their data. This study was performed in accordance with the Declaration of Helsinki and the ethics committee of Newcastle University.

5.2.2 Behavioural experiment

Stimuli

Note that this experimental procedure is exactly as in [179], Experiment 2. Auditory stimuli comprised amplitude-modulated tones with a carrier frequency of 250 Hz. Five tones were sinusoidally amplitude-modulated at a frequency of 2 Hz with modulation depths of: 52%, 44%, 36%, 28%, and 20% (*see* Figure 5.1). The tones were created in MATLAB 2012 with a 44.1

kHz sampling rate and saved as 1.5s stereo wav files. Sounds were presented at 75 dB SPL using Sennheiser HD380 pro headphones.

Visual stimuli were 1.5s videos of a size-modulated three-dimensional (3D) cuboid. In 3D Studio Max (Autodesk, Inc.) the "spherify" modifier was used to vary the size of a blue rectangular box of size 1.0 × 1.2 × 4.0 units (width × height × length). The modifier was varied sinusoidally at either 1 or 2 Hz and between values of 0.16 and 0.44 (where a value of 0 is the original cuboid and a value of 1 is a sphere, *see* Figure 5.2). The cuboid was rendered at an oblique angle against a uniform black background. The cuboid spanned a visual angle of 13.7° × 13.7° (300 × 300 pixels,*see* Figure 5.1). The videos were saved as Quicktime mov files (60 frames per second, H.264 compression) and were displayed using a Dell Ultrasharp 19" flat panel monitor (model: 1907FP) at a resolution of 1280 × 1024 pixels.

Procedure

The behavioural experiment was run in three blocks: one auditory-only and two auditory-visual. The auditory-only block was always completed first. Participants were asked to read a set of instructions and confirm their understanding before beginning the experiment. Participants then completed a short practice block, which was composed of 16 auditory-only trials and gave feedback via a short beep for an incorrect answer. Feedback was not given during the experiment. Throughout the experiment, participants were seated in a dimly lit sound attenuating chamber.

In auditory-only blocks, participants were asked to listen to two sounds, and ascertain whether they were same or different. Participants listened to the tones and determined whether they were the same or different by listening to the relative difference in amplitude modulation (i.e. did the volume of one tone fluctuate more than the other?). When the presented pairs were different, one of the tones had modulation depth 20% and the other tone had depth 28%, 36%, 44%, or 52%. The order of presentation of smaller and larger modulation depth was counterbalanced. When the presented pair was the same, the percentage depth was 20%, 28%, 36%, 44%, or 52%. This led to 5 percentage differences between stimuli: 0% (same), and 8%, 16%, 24%, or 32% (different). There were 16 trials per percentage difference in modulation depth. All 80 trials were completed in one block, with a self-timed break after 64 trials.

In auditory-visual blocks, participants were again asked to listen to two

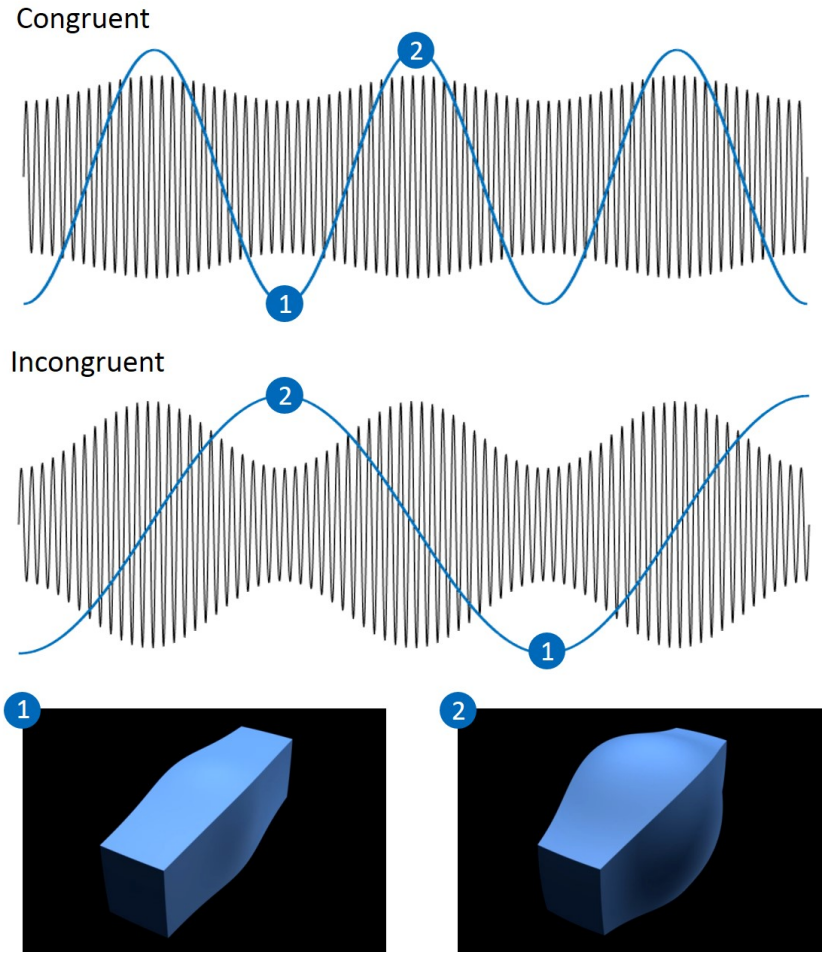

Figure 5.1: Example auditory and visual stimuli. The auditory tones were sinusoidally amplitude modulated at a frequency of 2Hz (black) and with a variable percentage depth (top: 20%, middle: 52%). The average amplitude was kept constant across modulation depths. The carrier frequency of the tones was 250 Hz. The shape size (blue) was modulated at either the same rate as the tone (2Hz, top), or a lower frequency (1Hz, middle) to yield congruent and incongruent combinations. Visual stimuli were blue cuboids, and a spherify modifier smoothly interpolated the vertices of the object into that of a sphere. The value of this modifier oscillated between low (0.16, *see* "1") and medium (0.44, *see* "2") values.

sounds and decide whether they were the same or different. A blue cuboid was presented alongside one of the amplitude-modulated tones, and a blue dot was presented alongside the other (*see* Figure 5.2). The presentation order of the shape and the dot was counterbalanced. In each trial, the cuboid was paired with either the greater or lesser modulated tone, and was temporally congruent with the tone (2Hz) or incongruent (1Hz). This led to 4 auditory-visual conditions: congruent greater modulation (Cong >), congruent lesser modulation (Cong <), incongruent greater modulation (Incong >), and incongruent lesser modulation (Incong <). There were 16 trials per dif-

ference level for each auditory-visual condition, leading to 320 trials. These were completed over two 160 trial blocks, with self-timed breaks after 64 and 128 trials. Participants were asked to continually attend the visual stimuli whilst discriminating the auditory stimuli.

In each trial, a fixation cross was presented for 0.5s (*see* Figure 5.2). An amplitude-modulated tone and a blue dot or cuboid was presented for 1.5s, followed by a blank screen for 1s. Another amplitude-modulated tone and blue dot or cuboid was presented for 1.5s, followed by a green dot which indicated that a response was required. The response dot was displayed until a response was given. A blank screen was displayed for 1 second before the next trial began.

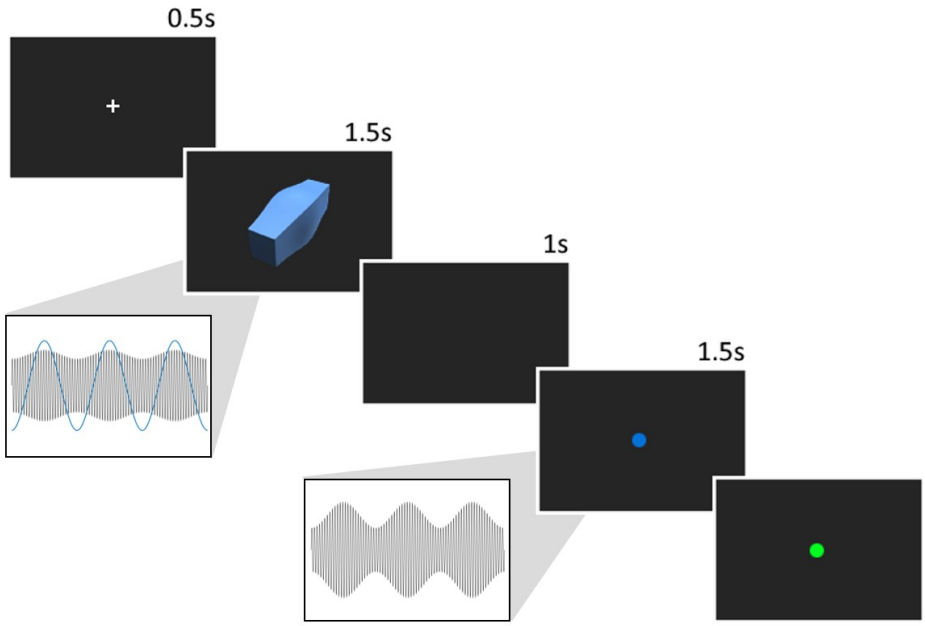

Figure 5.2: Trial time-course for an auditory-visual trial. In this example, a temporally congruent shape is paired with the tone with the lesser amplitude modulation. Participants heard the amplitude-modulated tones in two intervals and were asked whether the two sounds were the same or different. The shape modulation rate was either congruent (2 Hz, as in this example) or incongruent (1 Hz) to the amplitude modulation rate of the tone (2 Hz).

Analysis

For each participant and each condition, I calculated the average proportion of *different* responses for each modulation depth difference, and performed an initial ANOVA on the group. I then fit a cumulative Gaussian function to the resultant 5 datapoints.

During experimentation, an error occurred in which the order of presentation of the stimuli were not properly counterbalanced. For 1 participant, if the cuboid was to be presented alongside the greater amplitude-modulated tone, it was always presented during the first interval. Conversely, if the cuboid was to be paired with the lesser amplitude-modulated tone, it was always presented in the second interval.

Derived Behavioural Measures

In [179], the main effect of shape pairing occurred when the auditory-visual pair were temporally congruent. When the shape was paired with the lesser amplitude-modulated tone the overall proportion different response was reduced, and when paired with the greater amplitude-modulated tone, the overall proportion different response was increased. This resulted in a large separation between the proportion different for Cong $>$ and Cong $<$ conditions at each amplitude modulation difference (*see* [179]). A measure of integration strength was derived as the difference between 75% thresholds when the congruent shape was paired with the greater, and lesser modulated tone.

5.2.3 Spectroscopy

Procedure

After completing the behavioural study, participants were invited to take part in MRS spectroscopy providing that (a) their 75% threshold for amplitude modulation difference in all behavioural conditions could be accurately estimated, and (b) that their discrimination accuracy in the auditory-only task was above 70% for the largest amplitude modulation difference. In the scanner, soft pads were placed in the head coil to restrict movement, and participants were asked to lie as still as possible during the scans. A high-

resolution scan was initially performed, and MRS acquisition volumes were placed using the high-resolution image. Participants were allowed to hear music during the structural scan but not during spectroscopy acquisition.

Structural MRI

To perform volume segmentation of the MRS and accurately estimate GABA concentration, a high-resolution T1-weighted scan was performed. This scan consisted 150 slices, with repetition time (TR) = 9.6ms, echo time (TE) = 4.6ms, flip angle = 8°, field of view (FOV) 240mm × 240mm × 180mm, matrix size 208 × 208 pixels, voxel size 0.94mm × 0.94mm × 1.2mm, and took 5 minutes to complete.

MRS Acquisition

MRS was performed in two regions: the left superior temporal sulcus (STS), and right inferior frontal gyrus (IFG) (*see* Figure 5.3). The average voxel sizes in left STS or IPS were 24.8 mm ± 0.9 mm (anterior-posterior, mean ± standard error), 36.5 mm ± 2.2 mm (right-left) and 29.8 mm ± 1.8 mm (inferior-posterior).

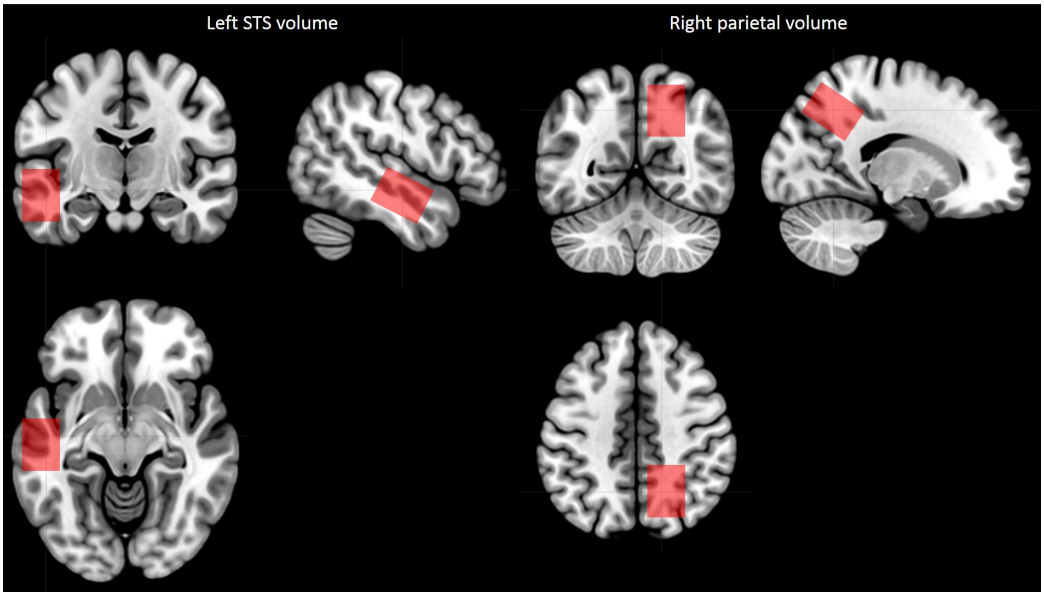

Figure 5.3: MRS acquisition volumes in the left temporal cortex (left) and right parietal cortex (right).

GABA+, which contains other macromolecules with the same resonance, was measured and the proportion of GABA within the GABA+ signal was subsequently estimated. GABA+ was measured with magnetic resonance spectroscopy (MRS) using the MEGA-PRESS method [132]. Specific parameters were: TR 2000ms, TE 68ms, 320 averages, acquisition bandwidth 1000Hz, and acquisition time 11 minutes. An inverse Gaussian editing pulse was applied at 1.9ppm during EDIT-ON scans and 7.5ppm during EDIT-OFF scans. The GABA+ signal at 3.02ppm was separated from the overlying creatine signal by subtracting the EDIT-OFF from the EDIT-ON spectra. Acquisitions took place on a Phillips Intera Acheiva 3 Tesla MRI scanner with an eight channel head coil. The MEGA-PRESS spectra were analysed using the GANNET pipeline [131] in MATLAB 2012, and Gaussian curves were fitted to the GABA+ spectral peaks [136]. If fitting could not be performed accurately (GANNET fit error $> 10\%$) the acquisition was withdrawn from correlation analyses. Non-water suppressed spectra were additionally obtained in both the left STS and right IPS (PRESS: TE 68ms; TR 2000 ms; 10 averages), and water amplitude was obtained from a Gaussian-Lorentzian fit to these spectra. Tissue fractions of grey matter, white matter, and cerebro-spinal fluid within the acquisition volumes were calculated using automated segmentation in the SPM8 package [168], using a volume mask generated from the acquisition volume. The final GABA measurement was estimated relative to water amplitude, tissue density, and the proportion of the GABA+ spectral peak thought to contain GABA discussed in [134, 187].

5.3 Results

5.3.1 Behaviour

The mean proportion of different responses are plotted in Figure 5.4. A 2 $\times$ 2 $\times$ 4 ANOVA (congruence $\times$ greater/lesser shape pairing $\times$ amplitude modulation difference, the 0% difference condition was not included) was carried out on the proportion of *different* responses. There was a significant effect of shape pairing ($F(1, 15) = 17.621, p < 0.000, \eta_p^2 = 0.718$), and a significant effect of amplitude modulation difference ($F(3, 45) = 116.833, p < 0.0005, \eta_p^2 = 0.886$). There was a significant interaction between shape pairing and congruence ($F(1, 15) = 4.669, p = 0.047, \eta_p^2 = 0.237$)

The mean 75% thresholds in each condition are shown in Figure 5.5. A 2×2 (congruence $\times$ greater/lesser shape pairing) repeated-measures ANOVA was performed on the 75% thresholds. There was a significant main effect of shape

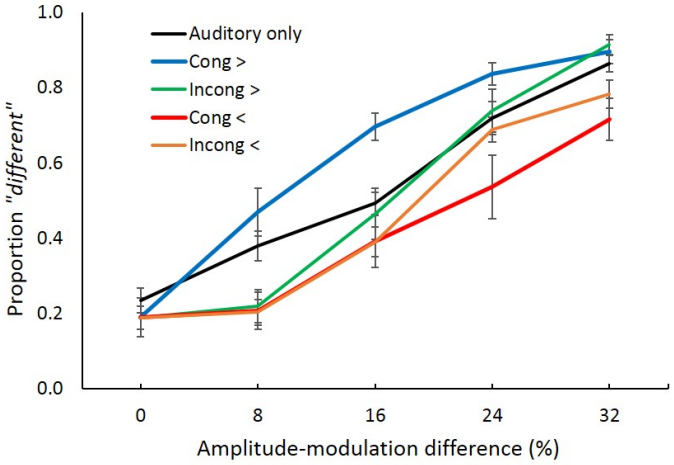

Figure 5.4: Proportion *"different"* responses as a function of the amplitude-modulation difference and the trial condition ($n = 16$). Error bars represent the standard error of the mean. The pattern of these results is comparable to those in [179].

pairing ($F(1, 15) = 12.884, p = 0.003, \eta_p^2 = 0.462$), but no effect of congruence and no interaction. Post-hoc pairwise comparisons were performed on the 75% threshold in the auditory-only condition against thresholds in each auditory-visual condition. Thresholds in the Cong > condition were significantly lower than in the auditory-only condition after Bonferroni correction ($t(15) = 2.293, p = 0.037$).

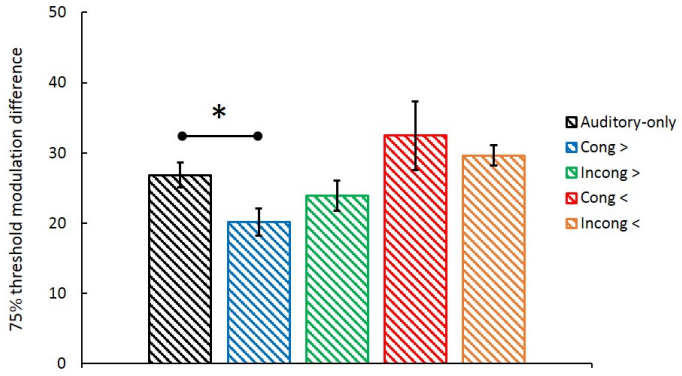

Figure 5.5: 75% thresholds for the proportion *different* responses in amplitude modulation difference for each experimental condition (n = 16). The threshold was significantly lower when a congruent shape was paired with the greater amplitude modulated tone than in any other condition (* $p < 0.05$).

5.3.2 Spectroscopy

The GANNET pipeline was ran on participants spectral data to estimate the volume of GABA within the parietal and temporal regions. No outliers were removed from the data prior to correlation analyses. Pearson correlation coefficients and uncorrected significance values are reported in Table 5.1. There was a correlation between parietal GABA concentration and 75% amplitude modulation threshold when the congruent shape was paired with the greater amplitude-modulated tone, and when an incongruent shape was paired with the greater amplitude-modulated tone. The derived measure of integration strength (difference between Cong > and Cong < thresholds) was also correlated with parietal GABA concentration (*see* Figure 5.6). However, no correlations were significant after correction for multiple comparisons.

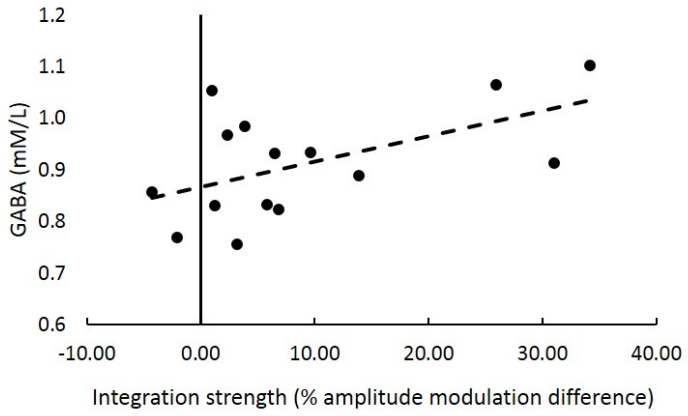

Figure 5.6: Parietal GABA concentration plotted against integration strength ($n = 16$). The correlation between these measurements was significant before correction for multiple comparisons $r = 0.512, p = 0.02$.

Threshold condition	Pearson's r	p
Auditory-only	-0.358	0.122
Cong >	-0.465	0.041
Incong >	-0.467	0.039
Cong <	0.151	0.329
Incong <	0.054	0.384
Integration strength	0.512	0.020

Table 5.1: Uncorrected correlation coefficients and significance values for parietal GABA and each behavioural measure ($n = 16$).

For temporal GABA, one participant had a fit error larger than 10%. Correlation analysis was conducted with the remaining 15 participants. There were no significant correlations in the temporal lobe (*see* Figure 5.7 Table 5.2).

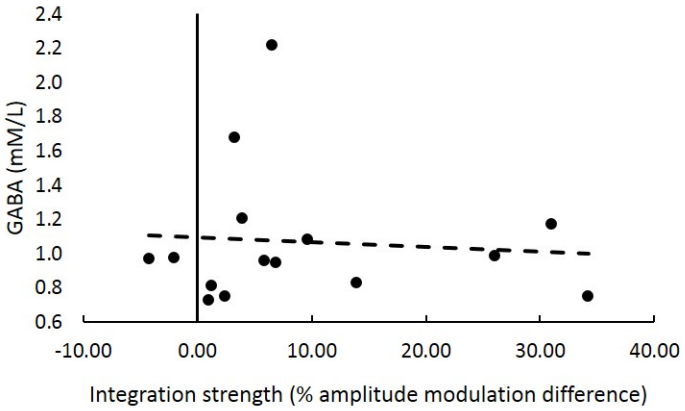

Figure 5.7: Temporal GABA concentration plotted against integration strength ($n = 15$).

Threshold condition	Pearson's r	p
Auditory-only	-0.014	0.961
Cong >	-0.109	0.700
Incong >	-0.079	0.780
Cong <	-0.192	0.815
Incong <	0.066	0.380
Integration strength	-0.082	0.771

Table 5.2: Uncorrected correlation coefficients and significance values for temporal GABA and each behavioural measure ($n = 15$).

5.4 Discussion

In this chapter, I hypothesised that GABAergic mechanisms, via inhibitive tuning or the generation of cortical gamma oscillations, facilitates multisensory integration in brain and behaviour. I measured the extent to which perception of an amplitude modulated tone was augmented by the presence of a temporally congruent or incongruent auditory stimulus. I then measured GABA in two putative multisensory regions, reported in [126] to demonstrate increased BOLD response to the same stimuli during congruence compared with incongruence. I found possible a correlation between the concentration of GABA in the parietal cortex and the strength of integration, which represented the extent to which auditory perception was augmented by the presence of a congruent visual stimulus. There were no correlations between GABA in the temporal lobe and any behavioural measure.

The posterior parietal cortex is thought to map the representation of space between separate sensory coordinate systems. It is possible that coordinate map tuning does occurs in the parietal lobe, facilitated by inhibition [182, 183]. This would led to increased multisensory responses to stimuli that are aligned in space. However, in my behavioural experiment, I did not vary

the spatial origin of the auditory and visual components. Inhibition in the parietal lobe may improve general perceptual cohesion between modalities via map tuning, but this interpretation requires further investigation. Future studies could explicitly manipulate spatial congruence, rather than temporal congruence, to correlate GABA with the behavioural outcome.

In the superior temporal sulcus (STS), GABA concentration was not correlated with any behavioural thresholds. This is perhaps surprising because there is evidence that the STS is sensitive to temporal congruence, which was manipulated in my behavioural experiment [14, 16, 103–108]. Temporal congruence was manipulated in my experiment, and the behavioural effect of shape pairing was absent for temporally incongruent stimuli. Given the lack of correlation, it is unlikely that the perception of temporal congruence was facilitated by GABA in my experiment. Low frequency amplitude modulation in my stimuli may have reduced reliance on gamma frequency oscillations for temporal synchrony, and by proxy reduced overall reliance on GABA concentration. In [146], GABA concentration in the STS predicted both gamma oscillatory power, and the extent to which participants perceived the sound induced flash illusion. The authors note that the stimuli presented during sound induced flash illusion paradigms consistently produce gamma band oscillations. It is possible that GABA has a general effect on multisensory timing, but this effect is reduced for stimuli with lower frequencies. More work is needed to ascertain the frequencies over which GABA concentration may influence multisensory timing and coordination.

While GABA may not be directly associated with the strength of multisensory integration, it is likely that inhibition plays a central role in both the spatial and temporal tuning of multisensory events.

Chapter 6

Conclusions

Multisensory integration allows us to combine and leverage signals from multiple modalities to reveal more about stimuli than could be revealed by the sum of the individual senses. A core problem that integrative mechanisms must overcome is to determine whether multisensory signals resulted from the same event in time. This is further complicated by the relative delay between arrival times of the physical signals, resulting from their different propagation speeds. For example, if an event emits auditory and visual signals at a distance of 20m, there is a 61ms delay between the arrival of light and sound to the observer. In this work, I investigated how the temporal asynchronies between audition and vision are managed in relation to attention, adaptation, and inhibition.

In Chapter 2, I preceded judgments about temporal order with regular and irregular temporal patterns. I hypothesised that a regular pattern would induce increased attention when stimuli could be reliably anticipated, as in dynamic attending theory (DAT), and that this increased attention would lead to a shorter temporal binding window (TBW). In one experiment, the TBW was shorter when the temporal order judgement (TOJ) was preceded by regular cues than irregular cues. I attributed the improved TOJ accuracy to increased attention that was induced by the regular pattern. However, I was unable to replicate this result in subsequent experiments, and sensitivity to temporal order in cued conditions may be a general consequence attentional cuing [32, 188]. It is possible that the regular cues were unreliable in inducing attentional entrainment, or that the irregular cues were unreliable in obstructing it. In this experiment, I was unable to verify that participants perceived the stimuli to be regular or irregular, nor could I verify that participants perceived any differences between the two conditions. Future work

on the effects of temporal context on the TBW could employ EEG to verify rhythmic entrainment at the neural level, or could embed TOJs within longer streams of regular or regular stimuli to ensure a rhythmic percept.

Recalibration studies demonstrate that exposure to temporally asynchronous stimuli causes a recalibration of the perceived simultaneity of events [38, 39]. Classical recalibration studies used extended "adaptation" periods to elicit this effect, but rapid recalibration studies demonstrate recalibration effects without an adaptation period. In Chapter 3, I measured perceptual simultaneity whilst participants were exposed to continuous, asynchronous stimuli. This allowed me to track the time-course of auditory-visual temporal adaptation. I demonstrated that adaptation occurs only for small auditory-leading asynchronies, but occurs over a large range of visual-leading asynchronies.

There is a general preference for visual-leading temporal asynchronies, with numerous studies reporting a non-zero, visual-leading PSS [27, 31, 32, 37, 83, 188–194]. One explanation for this effect is that light travels faster than sound, and so our senses are tuned to expect visual stimuli before auditory stimuli. This has further validity when considering that auditory stimuli are typically preceded by visual cues (i.e. the movement of a hammer before hitting a nail). Analyses of orofacial movements have shown that visual speech leads auditory speech by 100-300ms [64], though another study estimated the range to be between 30ms auditory-leading and 170ms visual-leading [195]. However, a within-participant comparison of the TBW and PSS demonstrated stronger asymmetries for beep-flash stimuli than for speech [31]. Another explanation for the visual-leading bias in synchrony judgements is that sensory signals take longer to reach the visual cortex than the auditory cortex [92]. Therefore if neural signals are to arrive simultaneously at some intersensory processing region, then light emitted by an event must reach an observer before sound (as is the natural case).

Classical recalibration studies have reported recalibration in both auditory-leading and visual-leading directions [38, 40, 41]. The symmetry effects evidenced in these studies are at odds with the findings reported in Chapter 3. This is surprising because each of these studies expose participants to auditory-visual asynchrony for extended periods. A key difference is that recalibration studies took SJ or TOJ measurements of a range of asynchronies *after* exposure. For classical recalibration, it is possible that exposure in the visual-leading direction altered the underlying mechanisms which determine temporal coincidence, causing a shift in perceptual response and measurable differences in synchrony and temporal order judgements. Conversely, extended exposure in the auditory-leading direction may simply increase the likelihood that test asynchronies are compared to the exposure asynchrony.

In this case, test asynchronies may be interpreted as synchronous simply because they are *less* auditory-leading than the exposure stimulus, but not because of any mechanistic changes. However, conclusions drawn via comparison between the present study and classical recalibration studies are indirect due to their methodical differences. Future work should present and measure responses to asynchrony during the adaptation phase as in Chapter 3, and measure the TBW and PSS with subsequent synchrony judgements. I could then determine whether auditory-leading adaptation is comparative only.

This thesis has focused on the mechanisms of integration that manage latency between audition and vision, but has not addressed other senses. Recalibration to temporal asynchrony is possible for auditory-tactile [45], and visual-tactile [44] stimuli, but not when the adaptation period is very short [194]. This has led investigators to conjecture that recalibration occurs at different time scales for audition and vision compared to audition and touch, or vision and touch [194]. Regarding Chapter 3, if haptic stimuli were used in place of auditory stimuli, future studies could investigate whether the asymmetry between the amount of adaptation for visual-leading and auditory-leading asynchronies resulted from the differing physical propagation speeds of sound and light, or is a property of the sensory adaptation in general.

In Chapter 4 I investigated the neural basis of the adaptation effects observed in Chapter 3. By manipulating the order of presentation of synchronous, auditory-leading, and visual-leading stimuli, I manipulated the state of adaptation in order to capture these states in fMRI. I additionally asked participants to report their perceptual simulaneity throughout each trial. This investigation reinforced the involvement of frontal and parietal regions in the perception of temporal asynchrony, in line with recent investigations [122, 126]. I identified the posterior cingulate cortex (PCC) as a possible monitor of and compensator for temporal asynchrony. The PCC has not been previously identified in multisensory literature, but is thought to have a general role in conflict monitoring and cognitive control [176]. Future work could focus on the PCC for temporal adaptation, and investigate functional connectivity in the region during periods of synchrony and asynchrony to interrogate its role as a conflict monitor.

It is perhaps surprising that activity in primary visual and auditory cortices was not increased during temporal synchrony, since this is often the case in similar fMRI studies [91, 104, 107, 117, 170, 196]. However, many other fMRI studies have not found evidence of increased activation in early sensory areas [16, 119, 122, 126]. Nonetheless, it is clear that early sensory regions have a large impact on multisensory processing. Numerous EEG studies have demonstrated how auditory stimuli can reset the phase of neural oscillations

in primary visual areas [78, 80, 82] and vice versa [81, 96]. Some studies additionally evidenced how these early sensory interactions can affect perception [77, 78, 81, 82, 197]. Given the wealth of evidence supporting the role of primary sensory areas in multisensery perception, it is unlikely that these areas are not involved in temporal adaptation. It is possible that processing in these regions depends upon neural oscillations and does not elicit an increased BOLD response. Future work should investigate the neural response to adaptation with EEG and determine the extent to which phase reset and rhythmic entrainment modulate the rate and amount of adaptation.

Activation in the PFC was increased during temporal asynchrony compared with temporal synchrony. This finding supports mounting evidence that the PFC region provides top-down modulatory feedback to sensory regions [122, 124, 126]. However, the resolution of temporal asynchrony may not require high-level feedback. Evidence that integrative mechanisms are occurring in low-level, primary sensory cortices is growing [80, 82, 92, 95, 97, 107, 198] (for a review, *see* [199]). Recent studies suggest that low level circuits can carry complex, contextual information [200]. This is important because fast resolution of temporal asynchrony allows for the grouping of low-level features that can be propagated upward to facilitate higher-order object representations [3, 201, 202]. It is possible that top-down feedback occurs when stimuli are presented over a longer duration, but not for transient stimuli. More work is needed to understand the role of top-down feedback in the resolution of temporal congruence.

In Chapter 5 I investigated the impact of inhibition on multisensory integration. I carried out a behavioural task known to elicit a strong integrative effect [179] and measured resting GABA concentration in the right parietal, and left temporal cortices. These results indicated a possible correlation between the strength of adaptation, and GABA concentration in the parietal lobe. I proposed that this correlation was a result of improved spatial tuning in the parietal lobe [183], leading to improved integration. However, this hypothesis is tentative because the spatial origin of the auditory and visual components of the stimulus were not manipulated in the behavioural task. Future work should attempt to correlate GABA with behavioural measures that are reliant on high spatial or temporal tuning.

I did not use naturalistic stimuli in any of the experiments presented in this thesis, so it is not clear how well these results generalise for stimuli with increased semantic content. However, the ecological validity of the stimuli used in Chapters 3, 4, & 5, which were continuous, and their amplitude modulation was either derived directly from speech or was modulated at rates predominating in the temporal envelope of speech. Sensitivity to tempo-

ral coherence between audition and vision is sensitive to stimulus type [31]. It would be useful to repeat the experiments presented in Chapter 3 with auditory-visual speech stimuli and compare the impact of semantic congruence on adaptation. I would expect adaptation effects to continuous speech stimuli to be large: speech is known to widen the window of integration [31], and contains both fine temporal structure [64] which will likely increase temporal correspondence between modalities [84].

The spatial alignment between senses is also context dependent. The "ventriloquism aftereffect" occurs when participants have been exposed to spatially mismatched acoustic and visual spaces [203, 204]. Similar to rapid recalibration studies, the ventriloquism aftereffect can also occur after a single brief exposure to auditory-visual stimuli that are mismatched in space [205]. It is thought that these operations are reliant upon spatial mappings between the senses, that are maintained in the parietal lobe [115] and superior colliculus (SC) [181]. In addition, spatial realignment may be optimised with inhibitory tuning mechanisms [182, 183]. In Chapter 5 I measured the concentration of an inhibitory neurotransmitter in the parietal region. The role of GABAergic inhibition in integration may be more directly associated with sensory tuning, rather than overall integration strength. Future work could investigate whether the strength and duration of the ventriloquism aftereffect is modulated by GABA mediated inhibition in the parietal region.

To conclude, I propose that low-level mechanisms govern the resolution of event timing between senses. These low-level mechanisms occur within and between primary sensory cortices, and calculate the statistical likelihood of the simultaniety of events. I propose that the parameters of this calculation are influenced by the temporal context of low-level stimulus features, such as frequency, rhythm, and stimulus envelope. These parameters are further influenced by top-down signals from frontal regions, which accumulate evidence over longer durations, as to whether or not multisensory signals are temporally linked. Overall, the work in this thesis supports the broader conclusion that the brain is a dynamic, context-driven, Bayesian estimator, which applies appropriate probability functions to the binding of event signals, and is able to do so at multiple timescales.

Bibliography

1. Feldman, J. The neural binding problem(s). *Cognitive Neurodynamics* **7,** 1–11. ISSN: 1871-4080. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3538094%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Feb. 2013).
2. Stein, B. E. & Stanford, T. R. Multisensory integration: current issues from the perspective of the single neuron. *Nature reviews. Neuroscience* **9,** 255–66. ISSN: 1471-0048. http://www.ncbi.nlm.nih.gov/pubmed/18354398 (Apr. 2008).
3. Bizley, J. K., Maddox, R. K. & Lee, A. K. Defining Auditory-Visual Objects: Behavioral Tests and Physiological Mechanisms. *Trends in Neurosciences* **39,** 74–85. ISSN: 1878108X. http://dx.doi.org/10.1016/j.tins.2015.12.007 (2016).
4. Bernstein, I. H., Clark, M. H. & Edelstein, B. A. Effects of an auditory signal on visual reaction time. eng. *Journal of experimental psychology* **80,** 567–569. ISSN: 0022-1015 (Print) (June 1969).
5. Gielen, S. C. A. M., Schmidt, r. A. & van den Heuvel, P. J. On the nature of intercensory facilitation of reaction time. *Perception & Psychophysics* **34,** 161–168. ISSN: 0031-5117 (1983).
6. Frassinetti, F., Bolognini, N. & Làdavas, E. Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research* **147,** 332–343. ISSN: 00144819 (2002).
7. Lovelace, C. T., Stein, B. E. & Wallace, M. T. An irrelevant light enhances auditory detection in humans: A psychophysical analysis of multisensory integration in stimulus detection. *Cognitive Brain Research* **17,** 447–453. ISSN: 09266410 (2003).
8. Meredith, A. M. & Stein, B. E. Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research* **365,** 350–354. ISSN: 0006-8993 (1986).
9. Sumby, W. H. & Pollack, I. Visual Contribution to Speech Intelligibility in Noise. *Journal of the Acoustical Society of America* **26,** 212–215 (1954).

10. Crosse, M. J., Di Liberto, G. M. & Lalor, E. C. Eye Can Hear Clearly Now: Inverse Effectiveness in Natural Audiovisual Speech Processing Relies on Long-Term Crossmodal Temporal Integration. *Journal of Neuroscience* **36,** 9888–9895. ISSN: 0270-6474. http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1396-16.2016 (2016).

11. Stevenson, R. a., Kim, S. & James, T. W. An additive-factors design to disambiguate neuronal and areal convergence: Measuring multisensory interactions between audio, visual, and haptic sensory streams using fMRI. *Experimental Brain Research* **198,** 183–194. ISSN: 00144819 (2009).

12. Stevenson, R. a. & James, T. W. Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *NeuroImage* **44,** 1210–1223. ISSN: 10538119. http://dx.doi.org/10.1016/j.neuroimage.2008.09.034 (2009).

13. Grant, K. W. & Seitz, P. P.-f. The use of visible speech cues for improving auditory detection of spoken sentences. *The Journal of the Acoustical Society of America* **108,** 1197–1208. ISSN: 00014966 (2000).

14. Calvert, G. a., Hansen, P. C., Iversen, S. D. & Brammer, M. J. Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *NeuroImage* **14,** 427–38. ISSN: 1053-8119. http://www.ncbi.nlm.nih.gov/pubmed/11467916 (Aug. 2001).

15. Meredith, M. a., Nemitz, J. W. & Stein, B. E. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of neuroscience* **7,** 3215–29. http://www.ncbi.nlm.nih.gov/pubmed/3668625 (1987).

16. Calvert, G. a., Campbell, R. & Brammer, M. J. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current biology : CB* **10,** 649–57. ISSN: 0960-9822. http://www.ncbi.nlm.nih.gov/pubmed/10837246 (June 2000).

17. Wallace, M. T., Wilkinson, L. K. & Stein, B. E. Representation and integration of multiple sensory inputs in primate superior colliculus. eng. *Journal of neurophysiology* **76,** 1246–1266. ISSN: 0022-3077 (Print) (Aug. 1996).

18. Lewald, J. & Guski, R. Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cognitive Brain Research* **16,** 468–478. ISSN: 09266410 (2003).

19. Chen, Y. C. & Spence, C. When hearing the bark helps to identify the dog: Semantically-congruent sounds modulate the identification of masked pictures. *Cognition* **114,** 389–404. ISSN: 00100277. http://dx.doi.org/10.1016/j.cognition.2009.10.012 (2010).

20. Laurienti, P. J., Kraft, R. a., Maldjian, J. a., Burdette, J. H. & Wallace, M. T. Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research* **158,** 405–414. ISSN: 00144819 (2004).

21. Van der Burg, E., Brederoo, S. G., Nieuwenstein, M. R., Theeuwes, J. & Olivers, C. N. L. Audiovisual semantic interference and attention: Evidence from the attentional blink paradigm. *Acta Psychologica* **134,** 198–205. ISSN: 00016918. http://dx.doi.org/10.1016/j.actpsy.2010.01.010 (2010).

22. Snodgrass, J. G. & Vanderwart, M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory* **6,** 174–215. ISSN: 00961515 (1980).

23. Welch, R. B. Chapter 15 Meaning, attention, and the "unity assumption" in the intersensory bias of spatial and temporal perceptions. *Advances in Psychology* **129,** 371–387. ISSN: 01664115 (1999).

24. Vroomen, J. & Keetels, M. Perception of intersensory synchrony : A tutorial review. *Attention, Perception & Psychophysics* **72,** 871–884 (2010).

25. Vatakis, A. & Spence, C. Crossmodal binding: evaluating the "unity assumption" using audiovisual speech stimuli. *Perception & psychophysics* **69,** 744–756. ISSN: 0031-5117 (2007).

26. Chen, Y. C. & Spence, C. Assessing the role of the 'unity assumption' on multisensory integration: A review. *Frontiers in Psychology* **8,** 1–22. ISSN: 16641078 (2017).

27. Poppel, E., Schill, K. & von Steinbuchel, N. Sensory integration within temporally neutral systems states: a hypothesis. *Naturwissenschaften* **77,** 89–91 (1990).

28. Lewis, R. & Noppeney, U. Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30,** 12329–12339. ISSN: 0270-6474 (2010).

29. Ogawa, A. & Macaluso, E. Audio-visual interactions for motion perception in depth modulate activity in visual area V3A. *NeuroImage* **71,** 158–167. ISSN: 10538119. http://dx.doi.org/10.1016/j.neuroimage.2013.01.012 (2013).

30. Maddox, R. K., Atilgan, H., Bizley, J. K. & Lee, A. K. Auditory selective attention is enhanced by a task-irrelevant temporally coherent visual stimulus in human listeners. *eLife* **4,** 1–11. ISSN: 2050-084X. http://elifesciences.org/lookup/doi/10.7554/eLife.04995 (2015).

31. Stevenson, R. a. & Wallace, M. T. Multisensory temporal integration: Task and stimulus dependencies. *Experimental Brain Research* **227,** 249–261. ISSN: 00144819 (2013).

32. Van Eijk, R. L. J., Kohlrausch, A., Juola, J. F. & van de Par, S. Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception & Psychophysics* **70,** 955–968. ISSN: 0031-5117. http://link.springer.com/10.3758/PP.70.6.955 (Aug. 2008).

33. Vatakis, A. & Spence, C. Audiovisual synchrony perception for music, speech, and object actions. *Brain Research* **1111,** 134–142. ISSN: 00068993 (2006).

34. Zampini, M., Guest, S., Shore, D. I. & Spence, C. Audio-visual simultaneity judgments. *Perception & psychophysics* **67,** 531–544. ISSN: 0031-5117 (2005).

35. Spence, C. & Squire, S. Multisensory integration: Maintaining the perception of synchrony. *Current Biology* **13,** 519–521. ISSN: 09609822 (2003).

36. Stevenson, R. a., Wilson, M. M., Powers, A. R. & Wallace, M. T. The effects of visual training on multisensory temporal processing. *Experimental Brain Research* **225,** 479–489. ISSN: 00144819 (2013).

37. Powers, A. R., Hillock, A. R. & Wallace, M. T. Perceptual training narrows the temporal window of multisensory binding. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29,** 12265–74. ISSN: 1529-2401. `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2771316%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract` (Sept. 2009).

38. Fujisaki, W., Shimojo, S., Kashino, M. & Nishida, S. Recalibration of audiovisual simultaneity. *Nature neuroscience* **7,** 773–8. ISSN: 1097-6256. `http://www.ncbi.nlm.nih.gov/pubmed/15195098` (July 2004).

39. Van der Burg, E., Alais, D. & Cass, J. Rapid recalibration to audiovisual asynchrony. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **33,** 14633–7. ISSN: 1529-2401. `http://www.ncbi.nlm.nih.gov/pubmed/24027264` (Sept. 2013).

40. Vroomen, J., Keetels, M., De Gelder, B. & Bertelson, P. Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research* **22,** 32–35. ISSN: 09266410 (2004).

41. Navarra, J. *et al.* Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Cognitive Brain Research* **25,** 499–507. ISSN: 09266410 (2005).

42. Vatakis, A., Navarra, J., Soto-Faraco, S. & Spence, C. Temporal recalibration during asynchronous audiovisual speech perception. *Experimental Brain Research* **181,** 173–181. ISSN: 00144819 (2007).

43. Roseboom, W. & Arnold, D. H. Twice Upon a Time: Multiple Concurrent Temporal Recalibrations of Audiovisual Speech. *Psychological science : a journal of the American Psychological Society / APS* **22,** 872–877. ISSN: 0956-7976. `http://pss.sagepub.com/lookup/doi/10.1177/0956797611413293` (2011).

44. Hanson, J. V. M., Heron, J. & Whitaker, D. Recalibration of perceived time across sensory modalities. *Experimental Brain Research* **185,** 347–352. ISSN: 00144819 (2008).

45. Navarra, J., Soto-Faraco, S. & Spence, C. Adaptation to audiotactile asynchrony. *Neuroscience Letters* **413,** 72–76. ISSN: 03043940 (2007).

46. Navarra, J., Hartcher-O'Brien, J., Piazza, E. & Spence, C. Adaptation to audiovisual asynchrony modulates the speeded detection of sound. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 9169–9173. ISSN: 0027-8424 (2009).

47. Harrar, V. & Harris, L. R. The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Experimental Brain Research* **186,** 517–524. ISSN: 00144819 (2008).

48. Van der Burg, E., Orchard-Mills, E. & Alais, D. Rapid temporal recalibration is unique to audiovisual stimuli. *Experimental Brain Research.* ISSN: 14321106 (2014).

49. Harvey, C., Van der Burg, E. & Alais, D. Rapid temporal recalibration occurs crossmodally without stimulus specificity but is absent unimodally. *Brain Research* **1585,** 1–11. ISSN: 00068993. `http://linkinghub.elsevier.com/retrieve/pii/S0006899314011032` (2014).

50. Van der Burg, E. & Goodbourn, P. Rapid , generalized adaptation to asynchronous audiovisual speech. *Proceedings of the Royal Society B* **282** (2015).

51. Kiyonaga, A., Scimeca, J. M., Bliss, D. P. & Whitney, D. Serial Dependence across Perception , Attention , and. *Trends in Cognitive Sciences* **21,** 493–497. ISSN: 1364-6613. `http://dx.doi.org/10.1016/j.tics.2017.04.011` (2017).

52. Fischer, J. & Whitney, D. Serial dependence in visual perception. *Nature Publishing Group* **17.** ISSN: 1097-6256 (2014).

53. Van der Burg, E., Alais, D. & Cass, J. Audiovisual temporal recalibration occurs independently at two different time scales. *Scientific Reports* **5,** 14526. ISSN: 2045-2322. `http://www.nature.com/articles/srep14526` (2015).

54. Yarrow, K., Jahn, N., Durant, S. & Arnold, D. H. Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Consciousness and Cognition* **20,** 1518–1531. ISSN: 10538100. `http://dx.doi.org/10.1016/j.concog.2011.07.003` (2011).

55. Alais, D., Ho, T., Han, S. & Burg, E. V. D. A matched comparison across three different sensory pairs of cross-modal temporal recalibration from sustained and transient adaptation. *i-Perception* **8,** 1–10. ISSN: 2041-6695 (2017).

56. Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A. & Schroeder, C. E. Neuronal Oscillations and Multisensory Interaction in Primary Auditory Cortex. *Neuron* **53,** 279–292. ISSN: 08966273 (2007).

57. Bendixen, A., Schröger, E. & Winkler, I. I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29,** 8447–8451. ISSN: 0270-6474 (2009).

58. Stefanics, G. *et al.* Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30,** 13578–13585. ISSN: 0270-6474 (2010).

59. Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I. & Schroeder, C. E. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science (New York, N.Y.)* **320,** 110–3. ISSN: 1095-9203. http://www.ncbi.nlm.nih.gov/pubmed/18388295 (Apr. 2008).

60. Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S. & Puce, A. Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences* **12,** 106–113. ISSN: 1364-6613. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3987824%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Mar. 2008).

61. Schroeder, C. E. & Lakatos, P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in neurosciences* **32,** 9–18. ISSN: 0166-2236. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2990947%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Jan. 2009).

62. Schroeder, C. E., Wilson, D. a., Radman, T., Scharfman, H. & Lakatos, P. Dynamics of Active Sensing and perceptual selection. *Current opinion in neurobiology* **20,** 172–6. ISSN: 1873-6882. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2963579%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Apr. 2010).

63. Van Atteveldt, N., Murray, M. M., Thut, G. & Schroeder, C. E. Multisensory integration: Flexible use of general operations. *Neuron* **81,** 1240–1253. ISSN: 10974199. http://dx.doi.org/10.1016/j.neuron.2014.02.044 (2014).

64. Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A. & Ghazanfar, A. a. The natural statistics of audiovisual speech. *PLoS Computational Biology* **5,** e1000436. ISSN: 1553-7358. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2700967%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (July 2009).

65. Plomp, R. in *Hearing: Physiological bases and psychophysics* (eds Kilnke, R. & R, H.) 270–276 (Springer-Verlag, 1983).

66. Jones, M. R. Time, Our Lost Dimension: Toward a New Theory of Perception, Attention, and Memory. *Psychological Review* **83,** 323–355. ISSN: 03622436 (1976).

67. Henry, M. J. & Herrmann, B. Low-Frequency Neural Oscillations Support Dynamic Attending in Temporal Context. *Timing and Time Perception* **2,** 62–86. ISSN: 22134468 (2014).

68. Large, E. W. & Jones, M. R. *The Dynamics of Attending: How People Track Time-Varying Events* 1999.

69. Barnes, R. & Jones, M. R. Expectancy, Attention, and Time. *Cognitive Psychology* **41,** 254–311. ISSN: 00100285 (2000).

70. Rohenkohl, G., Cravo, A. M., Wyart, V. & Nobre, A. C. Temporal Expectation Improves the Quality of Sensory Information. *Journal of Neuroscience* **32,** 8424–8428. ISSN: 0270-6474. http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0804-12.2012 (2012).

71. Drake, C. & Botte, M. C. Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. eng. *Perception & Psychophysics* **54,** 277–286. ISSN: 00315117 (Sept. 1993).

72. Miller, N. S. & McAuley, J. D. Tempo sensitivity in isochronous tone sequences: the multiple-look model revisited. eng. *Perception & psychophysics* **67,** 1150–1160. ISSN: 0031-5117 (Print) (Oct. 2005).

73. Bolger, D., Trost, W. & Schön, D. Rhythm implicitly affects temporal orienting of attention across modalities. *Acta Psychologica* **142,** 238–244. ISSN: 00016918. http://dx.doi.org/10.1016/j.actpsy.2012.11.012 (2013).

74. Escoffier, N., Sheng, D. Y. J. & Schirmer, A. Unattended musical beats enhance visual processing. *Acta Psychologica* **135,** 12–16. ISSN: 00016918 (2010).

75. Schulze, H. H. The detectability of local and global displacements in regular rhythmic patterns. *Psychological Research* **40,** 173–181. ISSN: 03400727 (1978).

76. Jones, M. R., Moynihan, H., MacKenzie, N. & Puente, J. Temporal Aspects of Stimulus-Driven Attending in Dynamic Arrays. *Psychological Science* **13,** 313–319. ISSN: 0956-7976. http://pss.sagepub.com/lookup/doi/10.1111/1467-9280.00458 (July 2002).

77. Romei, V., Gross, J. & Thut, G. Sounds reset rhythms of visual cortex and corresponding human visual perception. *Current Biology* **22,** 807–813. ISSN: 09609822. http://dx.doi.org/10.1016/j.cub.2012.03.025 (2012).

78. Fiebelkorn, I. C. *et al.* Ready, set, reset: stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **31,** 9971–9981. ISSN: 0270-6474 (2011).

79. Thorne, J. D. & Debener, S. Look now and hear what's coming: On the functional role of cross-modal phase reset. *Hearing Research* **307,** 144–152. ISSN: 03785955. http://dx.doi.org/10.1016/j.heares.2013.07.002 (2014).

80. Mercier, M. R. *et al.* Auditory-driven phase reset in visual cortex: human electrocorticography reveals mechanisms of early multisensory integration. *NeuroImage* **79,** 19–29. ISSN: 1095-9572. http://www.ncbi.nlm.nih.gov/pubmed/23624493 (Oct. 2013).

81. Thorne, J. D., De Vos, M., Viola, F. C. & Debener, S. Cross-modal phase reset predicts auditory task performance in humans. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **31,** 3853–61. ISSN: 1529-2401. http://www.ncbi.nlm.nih.gov/pubmed/21389240 (Mar. 2011).

82. Naue, N. *et al.* Auditory event-related response in visual cortex modulates subsequent visual responses in humans. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **31,** 7729–7736. ISSN: 1529-2401 (2011).

83. Arrighi, R., Alais, D. & Burr, D. Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *Journal of Vision* **6,** 6. ISSN: 1534-7362 (2006).

84. Denison, R. N., Driver, J. & Ruff, C. C. Temporal structure and complexity affect audio-visual correspondence detection. *Frontiers in Psychology* **3,** 1–12. ISSN: 16641078 (2013).

85. Mazza, V., Turatto, M., Rossi, M. & Umiltà, C. How automatic are audiovisual links in exogenous spatial attention? *Neuropsychologia* **45,** 514–522. ISSN: 00283932 (2007).

86. Santangelo, V., Van Der Lubbe, R. H., Olivetti Belardinelli, M. & Postma, A. Multisensory integration affects ERP components elicited by exogenous cues. *Experimental Brain Research* **185,** 269–277. ISSN: 00144819 (2008).

87. Spence, C. Crossmodal spatial attention. *Annals of the New York Academy of Sciences* **1191,** 182–200. ISSN: 17496632 (2010).

88. Mateeff, S., Hohnsbein, J. & Noack, T. Dynamic Visual Capture: Apparent Auditory Motion Induced by a Moving Visual Target. *Perception* **14,** 721–727. ISSN: 0301-0066. https://doi.org/10.1068/p140721 (Dec. 1985).

89. Bertelson, P. Ventriloquism: A Case of Crossmodal Perceptual Grouping. *Advances in Psychology* **129,** 347–362. ISSN: 0092-8674 (1999).

90. Alais, D. & Burr, D. Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology* **14,** 257–262. ISSN: 09609822 (2004).

91. Calvert, G. a. *et al.* Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport* **10,** 2619–2623. ISSN: 0959-4965 (1999).

92. Raij, T. *et al.* Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *European Journal of Neuroscience* **31,** 1772–1782. ISSN: 0953816X (2010).

93. Beer, A. L., Plank, T. & Greenlee, M. W. Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex. *Experimental brain research* **213,** 299–308. ISSN: 1432-1106. http://www.ncbi.nlm.nih.gov/pubmed/21573953 (Sept. 2011).

94. Beer, A. L., Plank, T., Meyer, G. & Greenlee, M. W. Combined diffusion-weighted and functional magnetic resonance imaging reveals a temporal-occipital network involved in auditory-visual object processing. *Frontiers in integrative neuroscience* **7,** 5. ISSN: 1662-5145. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3570774%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Jan. 2013).

95. Kayser, C., Petkov, C. I., Augath, M. & Logothetis, N. K. Integration of touch and sound in auditory cortex. *Neuron* **48,** 373–84. ISSN: 0896-6273. http://www.ncbi.nlm.nih.gov/pubmed/16242415 (Oct. 2005).

96. Kayser, C., Petkov, C. I. & Logothetis, N. K. Visual modulation of neurons in auditory cortex. *Cerebral cortex (New York, N.Y. : 1991)* **18,** 1560–74. ISSN: 1460-2199. http://www.ncbi.nlm.nih.gov/pubmed/18180245 (July 2008).

97. Kösem, A., Gramfort, A. & Van Wassenhove, V. Encoding of event timing in the phase of neural oscillations. *NeuroImage* **92,** 274–284. ISSN: 10959572. http://dx.doi.org/10.1016/j.neuroimage.2014.02.010 (2014).

98. Meredith, M. A. & Stein, B. E. Interactions Among Converging Sensory Inputs in the Superior Colliculus. *Science (New York, N.Y.)* **221,** 389–391 (1983).

99. Meredith, M. A. & Stein, E. Spatial Determinants of Multisensory Colliculus Neurons Integration in Cat Superior. *Journal of Neurophysiology* **75,** 1843–1857 (1996).

100. Stein, B. E., Huneycutt, W. S. & Meredith, M. A. Neurons and behavior: the same rules of multisensory integration apply. eng. *Brain research* **448,** 355–358. ISSN: 0006-8993 (Print) (May 1988).

101. Stein, B. E., Meredith, M. A., Huneycutt, W. S. & McDade, L. Behavioral Indices of Multisensory Integration: Orientation to Visual Cues is Affected by Auditory Stimuli. eng. *Journal of cognitive neuroscience* **1,** 12–24. ISSN: 0898-929X (Print) (1989).

102. Barraclough*, N. E. *et al.* Integration of Visual and Auditory Information by Superior Temporal Sulcus Neurons Responsive to the Sight of Actions. *Journal of Cognitive Neuroscience* **17,** 377–391. ISSN: 0898-929X. http://www.mitpressjournals.org/doi/10.1162/0898929053279586 (2005).

103. Wright, T. M., Pelphrey, K. a., Allison, T., McKeown, M. J. & McCarthy, G. Polysensory interactions along lateral temporal regions evoked by audio-visual speech. *Cerebral cortex (New York, N.Y. : 1991)* **13,** 1034–43. ISSN: 1047-3211. http://www.ncbi.nlm.nih.gov/pubmed/12967920 (Oct. 2003).

104. Beauchamp, M. S., Lee, K. E., Argall, B. D. & Martin, A. Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* **41,** 809–823. ISSN: 08966273 (2004).

105. Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H. & Martin, A. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nature neuroscience* **7,** 1190–2. ISSN: 1097-6256. http://www.ncbi.nlm.nih.gov/pubmed/15475952 (Nov. 2004).

106. Macaluso, E., George, N., Dolan, R., Spence, C. & Driver, J. Spatial and temporal factors during processing of audiovisual speech: a PET study. *NeuroImage* **21,** 725–32. ISSN: 1053-8119. http://www.ncbi.nlm.nih.gov/pubmed/14980575 (Feb. 2004).

107. Noesselt, T. *et al.* Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **27,** 11431–41. ISSN: 1529-2401. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2957075%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Oct. 2007).

108. Stevenson, R. a., Geoghegan, M. L. & James, T. W. Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. *Experimental Brain Research* **179,** 85–95. ISSN: 00144819 (2007).

109. Macaluso, E. & Driver, J. Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends in Neurosciences* **28,** 264–271. ISSN: 01662236 (2005).

110. Laurienti, P. J. *et al.* Cross-modal sensory processing in the anterior cingulate and medial prefrontal cortices. *Human Brain Mapping* **19,** 213–223. ISSN: 10659471 (2003).

111. Sugihara, T., Diltz, M. D., Averbeck, B. B. & Romanski, L. M. Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **26,** 11138–11147. ISSN: 0270-6474 (2006).

112. Sparks, D. L., Freedman, E. G., Chen, L. L. & Gandhi, N. J. Cortical and subcortical contributions to coordinated eye and head movements. *Vision Research* **41,** 3295–3305. ISSN: 00426989 (2001).

113. Bell, A. H., Meredith, M. A., Van Opstal, a. J. & Munoz, D. P. Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. *Journal of neurophysiology* **93,** 3659–3673. ISSN: 0022-3077 (2005).

114. Stricanne, B., Andersen, R. A. & Mazzoni, P. Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. eng. *Journal of neurophysiology* **76,** 2071–2076. ISSN: 0022-3077 (Print) (Sept. 1996).

115. Batista, A. P., Buneo, C. A., Snyder, L. H. & Andersen, R. A. Reach Plans in Eye-Centrered Coordinates. *Science* **285,** 257–260 (1999).

116. Andersen, R. A. & Cohen, Y. E. Reaches to sounds encoded in an eye-centered reference frame. *Neuron* **27,** 647–52. http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?db=m%7B%5C&%7Dform=6%7B%5C&%7Ddopt=r%7B%5C&%7Duid=11055445 (2000).

117. Werner, S. & Noppeney, U. Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30,** 2662–2675. ISSN: 0270-6474 (2010).

118. Ghazanfar, A. a., Maier, J. X., Hoffman, K. L. & Logothetis, N. K. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **25,** 5004–5012. ISSN: 0270-6474 (2005).

119. Vander Wyk, B. C. *et al.* Cortical integration of audio-visual speech and non-speech stimuli. *Brain and cognition* **74,** 97–106. ISSN: 1090-2147. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3869029%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Nov. 2010).

120. Naumer, M. J. *et al.* Investigating human audio-visual object perception with a combination of hypothesis-generating and hypothesis-testing fMRI analysis tools. *Experimental brain research* **213,** 309–20. ISSN: 1432-1106. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3155044%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Sept. 2011).

121. Sadaghiani, S., Maier, J. X. & Noppeney, U. Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29,** 6490–6499. ISSN: 0270-6474 (2009).

122. Noesselt, T., Bergmann, D., Heinze, H.-J., Münte, T. & Spence, C. Coding of multisensory temporal patterns in human superior temporal sulcus. *Frontiers in integrative neuroscience* **6,** 64. ISSN: 1662-5145. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3428803%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Jan. 2012).

123. Kim, H. *et al.* Brain Networks Engaged in Audiovisual Integration During Speech Perception Revealed by Persistent Homology-Based Network Filtration. ENG. *Brain connectivity.* ISSN: 2158-0022 (Electronic) (Mar. 2015).

124. Noppeney, U., Ostwald, D. & Werner, S. Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30,** 7434–7446. ISSN: 0270-6474 (2010).

125. Werner, S. & Noppeney, U. The contributions of transient and sustained response codes to audiovisual integration. *Cerebral Cortex* **21,** 920–931. ISSN: 10473211 (2011).

126. Laing, M., Rees, A. & Vuong, Q. C. Amplitude-modulated stimuli reveal auditory-visual interactions in brain activity and brain connectivity. *Frontiers in Psychology* **6,** 1–11 (2015).

127. Meredith, M. A., Wallace, M. T. & Stein, B. E. Visual, auditory and somatosensory convergence in output neurons of the cat superior colliculus: multisensory properties of the tecto-reticulo-spinal projection. eng. *Experimental brain research* **88,** 181–186. ISSN: 0014-4819 (Print) (1992).

128. Wolf, W., Hicks, T. & Albus, K. The contribution of GABA-mediated inhibitory mechanisms to visual response properties of neurons in the kitten's striate cortex. *The Journal of Neuroscience* **6,** 2779–2795. ISSN: 0270-6474 (1986).

129. Fuzessery, Z. M. & Hall, J. C. Role of GABA in shaping frequency tuning and creating FM sweep selectivity in the inferior colliculus. *Journal of Neurophysiology* **76,** 1059–1073. ISSN: 0022-3077 (1996).

130. Edden, R. A. E. & Barker, P. B. Spatial Effects in the Detection of Gamma-Aminobutyric Acid: Improved Sensitivity at High Fields Using Inner Volume Saturation. *Magnetic Resonance in Medicine* **58,** 1276–1282. ISSN: 07403194 (2007).

131. Edden, R. A. E., Puts, N. A. J., Harris, A. D., Barker, P. B. & Evans, C. J. Gannet: A batch-processing tool for the quantitative analysis of gamma-aminobutyric acid-edited MR spectroscopy spectra. *Journal of Magnetic Resonance Imaging* **40,** 1445–1452. ISSN: 15222586. arXiv: NIHMS150003 (2014).

132. Mescher, M., Merkle, H., Kirsch, J., Garwood, M. & Gruetter, R. Simultaneous in vivo spectral editing and water suppression. *NMR in biomedicine* **11,** 266–72. ISSN: 09523480. http://www.ncbi.nlm.nih.gov/pubmed/9802468 (Oct. 1998).

133. Gao, F. *et al.* Edited magnetic resonance spectroscopy detects an age-related decline in brain GABA levels. *NeuroImage* **78,** 75–82. ISSN: 1095-9572. http://www.ncbi.nlm.nih.gov/pubmed/23587685 (Sept. 2013).

134. Gao, F. *et al.* Decreased auditory GABA+ concentrations in presbycusis demonstrated by edited magnetic resonance spectroscopy. *NeuroImage* **106,** 311–6. ISSN: 1095-9572. http://www.ncbi.nlm.nih.gov/pubmed/25463460 (2015).

135. Sedley, W. *et al.* Human Auditory Cortex Neurochemistry Reflects the Presence and Severity of Tinnitus. *Neurobiology of Disease* **35,** 14822–14828 (2015).

136. Edden, R., Crocetti, D., Zhu, H., Gilbert, D. & Mostofsky, S. Reduced GABA concentration in attention-deficit/hyperactivity disorder: Discovery Service for Endeavour College of Natural Health Library. *Archives of General Psychiaty* **69,** 750–753 (2012).

137. Tayoshi, S. *et al.* GABA concentration in schizophrenia patients and the effects of antipsychotic medication: A proton magnetic resonance spectroscopy study. *Schizophrenia Research* **117,** 83–91. ISSN: 09209964 (2010).

138. Bartos, M., Vida, I. & Jonas, P. Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nature Reviews Neuroscience* **8,** 45–56. ISSN: 1471003X (2007).

139. Buzsáki, G. & Wang, X.-J. Mechanisms of Gamma Oscillations. *Annual Review of Neuroscience* **35,** 203–225. ISSN: 0147-006X. arXiv: NIHMS150003. http://www.annualreviews.org/doi/10.1146/annurev-neuro-062111-150444 (2012).

140. Sohal, V. S., Zhang, F., Yizhar, O. & Deisseroth, K. Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* **459,** 698–702. ISSN: 00280836 (2009).

141. Traub, R. D. *et al.* GABA-enhanced collective behavior in neuronal axons underlies persistent gamma-frequency oscillations. *Proceedings of the National Academy of Sciences* **100,** 11047–11052. ISSN: 0027-8424 (2003).

142. Edden, R. a. E., Muthukumaraswamy, S. D., Freeman, T. C. a. & Singh, K. D. Orientation discrimination performance is predicted by GABA concentration and gamma oscillation frequency in human primary visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29,** 15721–6. ISSN: 1529-2401. http://www.ncbi.nlm.nih.gov/pubmed/20016087 (Dec. 2009).

143. Puts, N. a. J., Edden, R. a. E., Evans, C. J., McGlone, F. & McGonigle, D. J. Regionally specific human GABA concentration correlates with tactile discrimination thresholds. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **31,** 16556–60. ISSN: 1529-2401. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3374929%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Nov. 2011).

144. Takeuchi, T., Yoshimoto, S., Shimada, Y., Kochiyama, T. & Kondo, H. M. Individual Differences in Visual Motion Perception and the Associated Excitatory and Inhibitory Neurotransmitter Concentrations in the Brain. *Journal of Vision* **16,** 42. ISSN: 1534-7362. http://jov.arvojournals.org/article.aspx?articleid=2492669 (2016).

145. Van Loon, A. M. *et al.* GABA shapes the dynamics of bistable perception. *Current Biology* **23,** 823–827. ISSN: 09609822. http://dx.doi.org/10.1016/j.cub.2013.03.067 (2013).

146. Balz, J. *et al.* GABA concentration in superior temporal sulcus predicts gamma power and perception in the sound-induced flash illusion. *NeuroImage* **125,** 724–730. ISSN: 10959572. http://dx.doi.org/10.1016/j.neuroimage.2015.10.087 (2016).

147. Hillock-Dunn, A. & Wallace, M. T. Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental science* **15,** 688–96. ISSN: 1467-7687. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4013750%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (Sept. 2012).

148. Brainard, D. H. The Psychophysics Toolbox. eng. *Spatial vision* **10,** 433–436. ISSN: 0169-1015 (Print) (1997).

149. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. eng. *Spatial vision* **10,** 437–442. ISSN: 0169-1015 (Print) (1997).

150. Grahn, J. a. See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research* **220,** 51–61. ISSN: 00144819 (2012).

151. Watson, A. B. & Pelli, D. G. QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics* **33,** 113–120 (1983).

152. King-Smith, P. E., Grigsby, S. S., Vingrys, a. J., Benes, S. C. & Supowit, a. Efficient and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation and practical implementation. *Vision research* **34,** 885–912. ISSN: 00426989 (1994).

153. Ward, L. M. & Mori, S. Attention cueing aids auditory intensity resolution. *Journal of the Acoustical Society of America* **100,** 1722–1727. ISSN: 0001-4966. http://www.ncbi.nlm.nih.gov/pubmed/8817898 (1996).

154. Posner, M. Orienting of attention. *The Quarterly journal of experimental psychology* **32,** 3–25. ISSN: 0033-555X. arXiv: arXiv:1011.1669v3 (1980).

155. Johnson, D. M. & Hafter, E. R. Uncertain-frequency detection: Cuing and condition of observation. *Perception & Psychophysics* **28,** 143–149. ISSN: 00315117 (1980).

156. Hawkins, H. L. *et al.* Visual Attention Modulates Signal Detectability. *Journal of Experimental Psychology: Human Perception and Performance* **16,** 802–811. ISSN: 00961523 (1990).

157. Powers, A. R., Hevey, M. a. & Wallace, M. T. Neural correlates of multisensory perceptual learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **32,** 6263–74. ISSN: 1529-2401. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3366559%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (May 2012).

158. IEEE Subcommittee on Subjective Measurements. IEEE Recommended Practices for Speech Quality Measurements. *IEEE Transactions on Audio and Electroacoustics* **17,** 227–246. ISSN: 0018-9278 (1969).

159. Mamassian, P. & Goutcher, R. Temporal dynamics in bistable perception. *Journal of vision* **5,** 361–375. ISSN: 1534-7362 (2005).

160. Grill-Spector, K. *et al.* Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* **24,** 187–203. ISSN: 08966273 (1999).

161. Henson, R. N. & Rugg, M. D. Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia* **41,** 263–270. ISSN: 00283932 (2003).

162. Stern, C. E. *et al.* The hippocampal formation participates in novel picture encoding: evidence from functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences* **93,** 8660–8665. ISSN: 0027-8424. http://www.pnas.org/cgi/doi/10.1073/pnas.93.16.8660 (1996).

163. Buckner, R. L. *et al.* Functional anatomical studies of explicit and implicit memory retrieval tasks. *Journal of Neuroscience* **15,** 12–29. ISSN: 02706474 (1995).

164. Desmond, E., Wagner, D. & Glover, H. Semantic Encoding and Retrieval in the Left Inferior Prefrontal Cortex: A Functional MRI Study of Task Difficulty and Process Specificity. *The Journal of neuroscience* **15,** 5870–5878. papers://717918dd-5917-4a8c-9b59-dee3afb2212f/Paper/p327 (1995).

165. Grill-Spector, K., Henson, R. & Martin, A. Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences* **10,** 14–23. ISSN: 13646613 (2006).

166. Schacter, D. L. & Buckner, R. L. Priming and the Brain. *Neuron* **20,** 185–195. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed%7B%5C&%7Dcmd=Retrieve%7B%5C&%7Ddopt=AbstractPlus%7B%5C&%7Dlist%7B%5C_%7Duids=9593113848698483 37related:EcUnNJ8oUA0J (1998).

167. Bushara, K. O., Grafman, J. & Hallett, M. Neural correlates of auditory-visual stimulus onset asynchrony detection. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **21,** 300–304. ISSN: 1529-2401 (2001).

168. Friston, K. J. *et al.* Statistical parametric maps in functional imaging: A general linear approach. *Human brain mapping.* **2,** 189–210. ISSN: 1065-9471 (1994).

169. Maldjian, J. A., Laurienti, P. J., Kraft, R. A. & Burdette, J. H. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage* **19,** 1233–1239. ISSN: 10538119 (2003).

170. Dhamala, M., Assisi, C. G., Jirsa, V. K., Steinberg, F. L. & Scott Kelso, J. a. Multisensory integration for timing engages different brain networks. *NeuroImage* **34,** 764–773. ISSN: 10538119 (2007).

171. Binder, M. Neural correlates of audiovisual temporal processing - Comparison of temporal order and simultaneity judgments. *Neuroscience* **300,** 432–447. ISSN: 18737544 (2015).

172. Van den Heuvel, M. P., Kahn, R. S., Goni, J. & Sporns, O. High-cost, high-capacity backbone for global brain communication. *Proceedings of the National Academy of Sciences* **109,** 11372–11377. ISSN: 0027-8424. http://www.pnas.org/cgi/doi/10.1073/pnas.1203593109 (2012).

173. Tomasi, D. & Volkow, N. D. Association between functional connectivity hubs and brain networks. *Cerebral Cortex* **21,** 2003–2013. ISSN: 10473211 (2011).

174. Carter, C. S. *et al.* Anterior Cingulate Cortex, Error Detection, and the Online Monitoring of Performance. *Science* **280,** 747–749. ISSN: 00368075 (1998).

175. Weissman, D. H., Gopalakrishnan, A., Hazlett, C. J. & Woldorff, M. G. Dorsal anterior cingulate cortex resolves conflict from distracting stimuli by boosting attention toward relevant events. *Cerebral Cortex* **15,** 229–237. ISSN: 10473211 (2005).

176. Leech, R., Kamourieh, S., Beckmann, C. F. & Sharp, D. J. Fractionating the Default Mode Network: Distinct Contributions of the Ventral and Dorsal Posterior Cingulate Cortex to Cognitive Control. *Journal of Neuroscience* **31,** 3217–3224. ISSN: 0270-6474. arXiv: 0507210v1 [arXiv:math]. http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5626-10.2011 (2011).

177. Hayden, B. Y., Smith, D. V. & Platt, M. L. Electrophysiological correlates of default-mode processing in macaque posterior cingulate cortex. *Proceedings of the National Academy of Sciences* **106,** 5948–5953. ISSN: 0027-8424. http://www.pnas.org/cgi/doi/10.1073/pnas.0812035106 (2009).

178. Van den Heuvel, M. P. & Sporns, O. Network hubs in the human brain. *Trends in Cognitive Sciences* **17,** 683–696. ISSN: 13646613. http://dx.doi.org/10.1016/j.tics.2013.09.012 (2013).

179. Vuong, Q. C., Laing, M., Prabhu, A., Tung, H. I. & Rees, A. Modulated stimuli demonstrate a non-Bayesian interaction between hearing and vision. *Submitted to Scientific Reports* (2019).

180. Shapley, R., Hawken, M., Ringach, D. L., by McLaughlin & their Collabo-. Dynamics of Orientation Selectivity in the Primary Visual Cortex and the Importance of Cortical Inhibition mental work and also on recent theoretical work and modeling. *Neuron* **38,** 689–699. https://ac.els-cdn.com/S0896627303003325/1-s2.0-S0896627303003325-main.pdf?%7B%5C_%7Dtid=66754c97-89ba-4e3c-99ac-7927774e2002%7B%5C&%7Dacdnat=1536167138%7B%5C_%7De2dc67b0bc776851af8053fef7105a12 (2003).

181. Jay, M. F. & Sparks, D. L. Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature* **309,** 345–347. ISSN: 00280836 (1984).

182. Davison, A. P. Learning Cross-Modal Spatial Transformations through Spike Timing-Dependent Plasticity. *Journal of Neuroscience* **26,** 5604–5615. ISSN: 0270-6474. http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5263-05.2006 (2006).

183. Friedel, P. & Van Hemmen, J. L. Inhibition, not excitation, is the key to multimodal sensory integration. *Biological Cybernetics* **98,** 597–618. ISSN: 03401200 (2008).

184. Lakatos, P. *et al.* An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of neurophysiology* **94,** 1904–1911. ISSN: 0022-3077 (2005).

185. Hasenstaub, A. *et al.* Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron* **47,** 423–435. ISSN: 08966273. arXiv: NIHMS150003 (2005).

186. Cardin, J. A. *et al.* Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature* **459,** 663–667. ISSN: 00280836. arXiv: NIHMS150003 (2009).

187. Mullins, P. G. *et al.* Current practice in the use of MEGA-PRESS spectroscopy for the detection of GABA. *NeuroImage* **86,** 43–52. ISSN: 10959572. arXiv: NIHMS150003. http://dx.doi.org/10.1016/j.neuroimage.2012.12.004 (2014).

188. Ikumi, N. & Soto-Faraco, S. Selective attention modulates the direction of audio-visual temporal recalibration. *PLoS ONE* **9.** ISSN: 19326203 (2014).

189. Mégevand, P., Molholm, S., Nayak, A. & Foxe, J. J. Recalibration of the Multisensory Temporal Window of Integration Results from Changing Task Demands. *PLoS ONE* **8.** ISSN: 19326203 (2013).

190. Van Wassenhove, V., Grant, K. W. & Poeppel, D. Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* **45,** 598–607. ISSN: 0028-3932. http://www.ncbi.nlm.nih.gov/pubmed/16530232 (Feb. 2007).

191. Stevenson, R. a., Altieri, N. a., Kim, S., Pisoni, D. B. & James, T. W. Neural processing of asynchronous audiovisual speech perception. *NeuroImage* **49,** 3308–3318. ISSN: 10538119. http://dx.doi.org/10.1016/j.neuroimage.2009.12.001 (2010).

192. Vroomen, J. & Stekelenburg, J. J. Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. *Journal of cognitive neuroscience* **22,** 1583–1596. ISSN: 0898-929X (2010).

193. Navarra, J., Fernández-Prieto, I. & Garcia-Morera, J. Realigning thunder and lightning: Temporal adaptation to spatiotemporally distant events. *PLoS ONE* **8,** 1–10. ISSN: 19326203 (2013).

194. Burg, E. V. D., Orchard-Mills, E. & Alais, D. Rapid temporal recalibration is unique to audiovisual stimuli: no effects for audiotactile or visuotactile stimuli. *IMRF2014(oral)* **8,** 47. ISSN: 14321106 (2014).

195. Schwartz, J. L. & Savariaux, C. No, There Is No 150 ms Lead of Visual Speech on Auditory Speech, but a Range of Audiovisual Asynchronies Varying from Small Audio Lead to Large Audio Lag. *PLoS Computational Biology* **10,** e1003743. ISSN: 1553-7358. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4117430%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (July 2014).

196. Van Atteveldt, N. M., Formisano, E., Blomert, L. & Goebel, R. The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cerebral Cortex* **17,** 962–974. ISSN: 10473211 (2007).

197. Cappe, C., Thut, G., Romei, V. & Murray, M. M. Auditory-visual multisensory interactions in humans: timing, topography, directionality, and sources. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30,** 12572–80. ISSN: 1529-2401. http://www.ncbi.nlm.nih.gov/pubmed/20861363 (Sept. 2010).

198. Macaluso, E. Multisensory processing in sensory-specific cortical areas. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* **12,** 327–338. ISSN: 1073-8584 (2006).

199. Driver, J. & Noesselt, T. Multisensory Interplay Reveals Crossmodal Influences on 'Sensory-Specific' Brain Regions, Neural Responses, and Judgments. *Neuron* **57,** 11–23. ISSN: 08966273. http://dx.doi.org/10.1016/j.neuron.2007.12.013 (2008).

200. Van Atteveldt, N. M., Peterson, B. S. & Schroeder, C. E. Contextual control of audiovisual integration in low-level sensory cortices. *Human Brain Mapping* **35,** 2394–2411. ISSN: 10970193 (2014).

201. Atilgan, H. *et al.* Integration of Visual Information in Auditory Cortex Promotes Auditory Scene Analysis through Multisensory Binding. *Neuron* **97,** 640–655.e4. ISSN: 10974199. https://doi.org/10.1016/j.neuron.2017.12.034 (2018).

202. Kourtzi, Z. & Connor, C. E. Neural Representations for Object Perception: Structure, Category, and Adaptive Coding. *Annual Review of Neuroscience* **34,** 45–67. ISSN: 0147-006X (2011).

203. Bertelson, P., Frissen, I., Vroomen, J. & De Gelder, B. The aftereffects of ventriloquism: Patterns of spatial generalization. *Perception and Psychophysics* **68,** 428–436. ISSN: 00315117 (2006).

204. Chen, L. & Vroomen, J. Intersensory binding across space and time: a tutorial review. *Attention, perception & psychophysics* **75,** 790–811. ISSN: 1943-393X. http://www.ncbi.nlm.nih.gov/pubmed/23709064 (2013).

205. Recanzone, G. H. Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proceedings of the National Academy of Sciences of the United States of America* **95,** 869–75. ISSN: 0027-8424. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=33810%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract (1998).