Ipser, A. (2014). Individual differences in audiovisual integration and timing. (Unpublished Doctoral thesis, City University London)

**Original citation**: Ipser, A. (2014). Individual differences in audiovisual integration and timing. (Unpublished Doctoral thesis, City University London)

# INDIVIDUAL DIFFERENCES IN AUDIOVISUAL INTEGRATION AND TIMING

ALBERTA IPSER

A THESIS SUBMITTED TO THE
DEPARTMENT OF PSYCHOLOGY
CITY UNIVERSITY LONDON
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

May 2014

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEDGEMENTS

I would like to thank the Psychology department at City University London, who, through their award of studentship made the work in this thesis possible.

I am grateful to my family for their support throughout my years of higher education.

I will always be thankful to have had the opportunity to carry out my research degree alongside amazing past and present PhD students in the Psychology Department at City University , whose friendship and moral support has kept my sanity in check.

I would like to thank the CNRU lab for providing a motivating and friendly work environment.

I am grateful to those whom I have had the opportunity to collaborate with and the skills I have learned from them, in particular Ryota Kanai and Marinella Cappelletti.

I would like to thank Dr Richard Cook for giving me the opportunity to further my research career whilst finishing this thesis.

Finally, I would like to express the deepest gratitude to my supervisor, Dr Elliot Freeman for the amazing guidance, encouragement and inspiration he provided me with throughout my PhD training and for his ceaseless amount of patience and trust in my abilities.

# DECLARATION

I grant powers of discretion to the University Librarian to copy this thesis in part or in whole without further reference to me. This covers only single copies for study purposes. The contents of this thesis are subject to normal conditions of acknowledgement. The content of this thesis is my own work.

# ABSTRACT

Sight and sound are processed in different parts of the brain and at different times, creating discrepancies between the relative arrival time of auditory and visual information at primary and multisensory cortices. Despite this, a commonly accepted view is that the brain strives for and achieves temporal unity across different sensory modalities. Using individual differences in subjective synchrony and audiovisual temporal processing, this thesis examines whether audiovisual synchronisation across different audiovisual processes is ever actually achieved and whether the timing of multisensory events is supported by unified or disparate mechanisms. Chapter 2 examines whether estimates of subjective synchrony across audiovisual integration and explicit temporal judgements are consistent within and between individuals. This chapter finds remarkable disunity in subjective audiovisual timing within individuals, characterised by negatively correlated estimates of perceptual asynchrony across tasks, which challenge existing accounts of how the nervous system maintains temporal coherence. Instead, a new theory of temporal renormalisation is proposed, whereby the relative timing of audiovisual signals within different mechanisms is perceived relative to the average timing across mechanisms. Chapter 3 reveals that individual differences in audiovisual synchronisation across different tasks are reflected in the structural variability of distinct brain clusters, suggesting that audiovisual relative timing is processed by multiple task-specific temporal mechanisms, whose performance is supported by distinct neural substrates. Chapter 4 explores the possibility that these perceptual mechanisms might contribute to reading ability, which is audiovisual in nature. Aspects of audiovisual temporal processing are found to be impaired in dyslexia and linearly related to reading ability. Altogether this thesis provides novel contributions to the understanding of the underlying mechanisms of audiovisual temporal processing as well as of its relationship to higher cognitive functions.

# Abbreviations

**AV:** Audiovisual

**ePSS**: Point of explicit subjective synchrony

**iPSS**: Point of implicit subjective synchrony

**JND**: Just-Noticeable Difference

**McG**: McGurk

**MFG**: Middle Frontal Gyrus

**MTG**: Middle temporal gyrus

**PSS**: Point of subjective synchrony

**SB**: Stream-Bounce

**SD**: Standard deviation (here, of cumulative Gaussian)

**SJ**: Synchrony Judgements

**STG**: Superior temporal Gyrus

**STS**: Superior temporal Cortex

**TMS:** Transcranial Magnetic Stimulation

**TOJ**: Temporal order judgements

**VBM**: Voxel-Based Morphology

**Win**: Window of synchrony/ AV integration

**/ /**: slashes indicate auditory speech sound of graphemes inside them

**[ ]**: brackets indicate visual lip-movements of graphemes inside them

# GLOSSARY

***Explicit point of subjective synchrony (ePSS):*** The audiovisual asynchrony at which participants perceive audiovisual stimuli as synchronous, which can be measured using temporal order judgements or synchrony judgements.

***Implicit point of subjective synchrony (iPSS):*** The audiovisual asynchrony at which participants integrate audiovisual information most often, which can be measured using audiovisual illusions as a function of audiovisual asynchrony.

***Temporal order judgements (TOJ):*** Temporal judgement task in which participants have to indicate the temporal order of audiovisual events presented at various audiovisual asynchronies. In the current thesis, participants were asked to indicate whether the sound came first or second.

***Synchrony judgements (SJ):*** Temporal judgement task in which participants have to indicate whether audiovisual stimuli pairs presented at various audiovisual asynchronies are synchronous or asynchronous.

***Temporal window of synchrony (SJ Win):*** Represents the range of audiovisual asynchronies within which individuals explicitly perceive audiovisual information to be synchronous. In this thesis, the temporal Window of audiovisual synchrony is derived from SJ data, and estimated by calculating the difference in means of the two cumulative Gaussians fitted to proportion of ‘synchronous’ responses plotted as a function of auditory lag. Higher measures represent smaller temporal specificity of audiovisual synchrony perception.

***Temporal window of audiovisual integration (McG/Stream-bounce Win):*** Represents the range of audiovisual asynchronies within which individuals integrate audiovisual information. In this thesis, the temporal Window of audiovisual integration is estimated by calculating the differences in means of the two cumulative Gaussians fitted to proportion of audiovisual integration plotted as a function of auditory lag. Higher measures represent smaller temporal specificity of audiovisual integration.

***Explicit discrimination of audiovisual synchrony:*** Represents how abruptly an observer switches from a synchronous response to an asynchronous response in SJ and from a sound first to a sound second response in TOJ, as a function of audiovisual asynchrony. This aspect of explicit temporal processing performance is estimated using the standard deviation of the single cumulative Gaussian fitted to TOJs, and the mean standard deviation of the two cumulative Gaussians fitted to SJs. Higher measures represent poorer sensitivity.

***Implicit discrimination of audiovisual synchrony:*** Represents how abruptly an observers' rate of audiovisual integration decreases as a function of audiovisual asynchrony. This aspect of implicit temporal processing performance is estimated using the average standard deviation of the two cumulative Gaussians fitted to audiovisual integration data. Higher measures represent poorer sensitivity.

***Implicit temporal processing:*** Processing of audiovisual relative timing which is not necessarily consciously accessed by the observer, for example during audiovisual integration.

***Explicit temporal processing:*** Process which underlies the conscious access to the relative timing of audiovisual information, for example during SJs and TOJs.

***Just noticeable difference:*** The smallest audiovisual asynchrony which can be detected reliably by an observer. Higher measures represent poorer sensitivity.

***Dyslexia:*** A neurobiological condition characterised by reading problems such as recognition, spelling and decoding of words, despite otherwise typical reading instruction and educational or professional attainment.

***Voxel-Based Morphometry (VBM):*** Voxel-wise analysis of the local concentration of grey matter, which can be used in group comparisons as well as correlational investigations. In this thesis VBM was used to correlate individual differences in behavioural measures with individual differences of local grey matter volume.

***Temporal recalibration:*** A shift in the perception of audiovisual relative timing or synchrony after prolonged exposure to a constant audiovisual asynchrony, represented by a shift in the PSS. This shift occurs in the direction of the audiovisual asynchrony observers are exposed to e.g. after exposure to a constant and prolonged auditory lag, auditory-lagging audiovisual stimuli will be more likely to be perceived as synchronous. Also referred to as temporal adaptation, recalibration is believed to maintain temporal coherence across sensory modalities.

***Renormalisation:*** audiovisual synchrony estimates within localised, task-specific temporal mechanisms are assessed in relation to the average asynchrony across all temporal mechanisms, leading to an antagonistic relationship between estimates of subjective timing.

***McGurk illusion:*** audiovisual speech illusion whereby auditory perception of a speech sound is affected by incongruent visual information. For example, observers will often hear /da/ when viewing lip-movements uttering [ga] presented together with the speech sound /ba/.

***Stream-Bounce Illusion:*** audiovisual non-speech illusion whereby the perceived trajectory of a pair of visual stimuli is affected by the occurence of a single beep. Two identical disks are displayed at each corner of a computer display, after which each one begins to move along a downward diagonal trajectory, crossing over each other at the midpoint of the display. Presenting a beep at or near the collision point increases likelihood of perceiving the disks colliding and changing trajectory.

# 1 CHAPTER 1: INTRODUCTION

Our senses are constantly flooded with sensory information belonging to different modalities, some of which originate from a common source and others which do not. For accurate and reliable interpretation of the environment, sensory streams which share various cues which transcend vision and audition are integrated into multisensory events. Temporal coincidence is one such cue and is viewed to be probably the most important amodal factor in audiovisual (AV) integration (e.g. Keetels & Vroomen, 2012; Spence & Squire, 2003). Presumably because in the natural world events which occur at the same time often originate from the same source, streams of AV information are more likely on average to be integrated when they are synchronous rather than asynchronous (Lewald & Guski, 2003; Meredith & Nemitz, 1987; van Atteveldt, Formisano, Blomert, & Goebel, 2007; van Wassenhove, Grant, & Poeppel, 2007). For example, AV synchrony is such a compelling cue, that when presented at the same time, visual and auditory information originating from different sources is bound together and perceived to originate from a single location (Slutsky & Recanzone, 2001).

The task of audiovisual synchronisation is unlikely to be computationally straightforward. For example, at approximately 300 million metres per second, the speed of light is much faster than the 340 metres per second at which sound travels through air. As a result, the visual component of an audiovisual event will always reach an observer before the auditory component (King, 2005; Spence & Squire, 2003). The difference between the arrival times of auditory and visual stimuli at sensory receptors will also increase as a function of the observer-stimulus distance (Sugita & Suzuki, 2003). Internally however, light energy has to be converted into chemically mediated nervous impulses and then in turn, into electrical signals. This process takes around 50ms longer than acoustic transduction, in which mechanical energy in the form of vibrations physically opens ion channels in the cochlear nerve, producing action potentials. Neural transmission time within the visual system is also longer compared to the auditory system, further increasing the temporal disparity between the time auditory and visual signals reach their relevant destinations (King, 2005). When travelling from a distance of 15m, known as the horizon of simultaneity, auditory and visual signals are said to arrive at primary sensory cortices at the same time (Pöppel, 1988). At distances shorter than the horizon of simultaneity, visual signals are said to lag the auditory whereas at distances beyond it they are said to lead. Auditory and visual signals have to then in turn converge on various multimodal sites in order to be processed as multisensory events (Benoit, Raij, Lin, Jääskeläinen, & Stufflebeam, 2010; Bertini, Leo, AVenanti, & Làdavas, 2010; Noesselt, Bergmann, Heinze, Münte, & Spence, 2012; Sekiyama, 2003). These cortical destinations will likely depend on the type of information being processed and the task being performed. The relative arrival time of auditory and visual signals will vary as a function of the cortical site at which they converge. Together, this is likely to lead to a cacophony of estimates of the relative timing of audiovisual information across different mechanisms.

The following introduction will begin with a description of the way in which audiovisual synchrony perception is measured using tasks which require observers to explicitly access information about the relative timing of AV information. This will be followed by an overview of two AV integration tasks which can be used to measure synchrony perception implicitly, without the need for the observer to make explicit judgements on the relative timing of AV information. The introduction will then discuss the potential benefits of examining behavioural individual differences in AV temporal processing as well as individual variability in brain structure in order to inform on the potential underlying mechanisms of explicit temporal judgements and AV integration.

# 1.1 MEASURING AUDIOVISUAL SYNCHRONY PERCEPTION

## 1.1.1 SYNCHRONY JUDGEMENTS (SJ) AND TEMPORAL ORDER JUDGEMENTS (TOJ)

Conscious or explicit access to the relative timing of AV events can be particularly useful in everyday life. For example, animators who work in the film industry are required to synchronize the movements of animated characters to the speech and sounds previously recorded by voice and sound effects artists. Television or film editors also have the task of synchronising pre-recorded audio with film clips to ensure that any delays are unnoticeable (Advanced Television Systems Committee, 2003). Although a proportion of such tasks are performed using automated software, the final outcome also contains human judgement. In the laboratory, subjective perception of AV synchrony (e.g. Navarra, Alsius, Velasco, Soto-Faraco, & Spence, 2010; Vatakis & Spence, 2006; Yarrow, Jahn, Durant, & Arnold, 2011) and sensitivity to the relative timing of crossmodal events (Keetels & Vroomen, 2005; Nicol & Shore, 2007; Roseboom, Nishida, Fujisaki, & Arnold, 2011; Vatakis & Spence, 2006) can be measured using simultaneity judgements (SJ) or temporal order judgements (TOJ), both of which require explicit access to the relative timing of AV events (Martin, Giersch, Huron, & van Wassenhove, 2012).

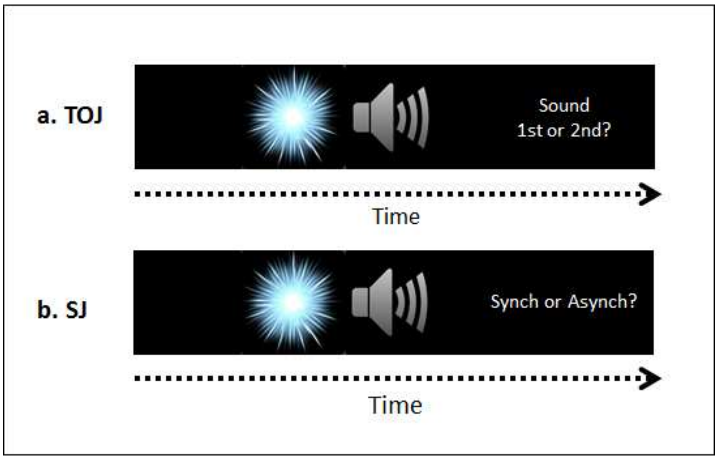


*Figure 1-1: Typical **a.** TOJ trial and **b.** SJ trial in which the light precedes the sound.*

In both SJ and TOJ paradigms (see Figure 1.1 for typical trials), audiovisual stimuli pairs are presented to participants at various stimulus onset asynchronies (SOA). For SJs individuals are asked to determine whether the auditory and visual streams were presented synchronously or asynchronously. The proportion of 'synchronous' responses is then plotted as a function of auditory lead or lag. As shown in Figure 1.2, this forms a bell shaped curve from which the asynchrony corresponding to the highest proportion of 'simultaneous' responses can be read. This asynchrony represents the observer's point of subjective simultaneity (PSS). For TOJs, individuals are asked to determine the temporal order in which the auditory and visual components of the AV stimulus were presented. The Proportion of 'sound first' (or light first) responses is then plotted as a function of auditory lead or lag.

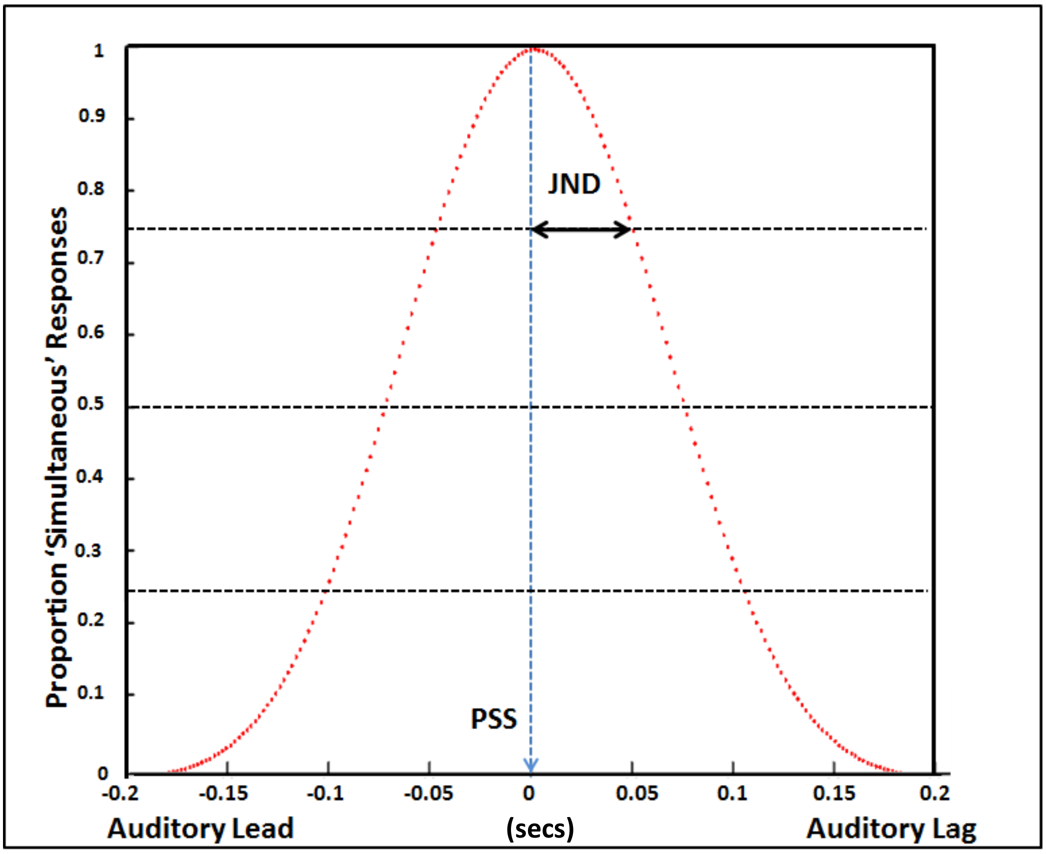


*Figure 1-2: Hypothetical data of a synchrony judgement (SJ) task, in which participants judge whether the sound and light occurred synchronously or asynchronously. The point of subjective synchrony (PSS) is value on the x axis which corresponds to the peak of the psychometric curve, in other words the highest proportion of 'synchronous' responses. In this scenario, the PSS is observed when the auditory and visual information is presented synchrony. The just-noticeable difference (JND) is typically derived by halving the distance between the two SOAs at which the participant responded 'synchronous' 75% of the time. Here, the JND is roughly 50ms.*

As shown in Figure 1.3 below, this forms an ‘S’ shaped curve to which a cumulative Gaussian function is normally fitted. The asynchrony at which participants respond to chance (in other words at which equal proportions of sound first and sound second responses are observed) represents the participant’s PSS.

Another measure that is derived from TOJ and SJ curves is the Just Noticeable Difference (JND). This measure represents the smallest asynchrony at which an observer can reliably judge whether the visual and auditory components were presented synchronously or the temporal order in which an AV stimulus was presented. For SJs, this measure is usually derived by halving the distance between the two SOAs at which the participant responded ‘synchronous’ 75% of the time ( (Vroomen & Keetels, 2010) (see Figure 1-2 on the previous page).

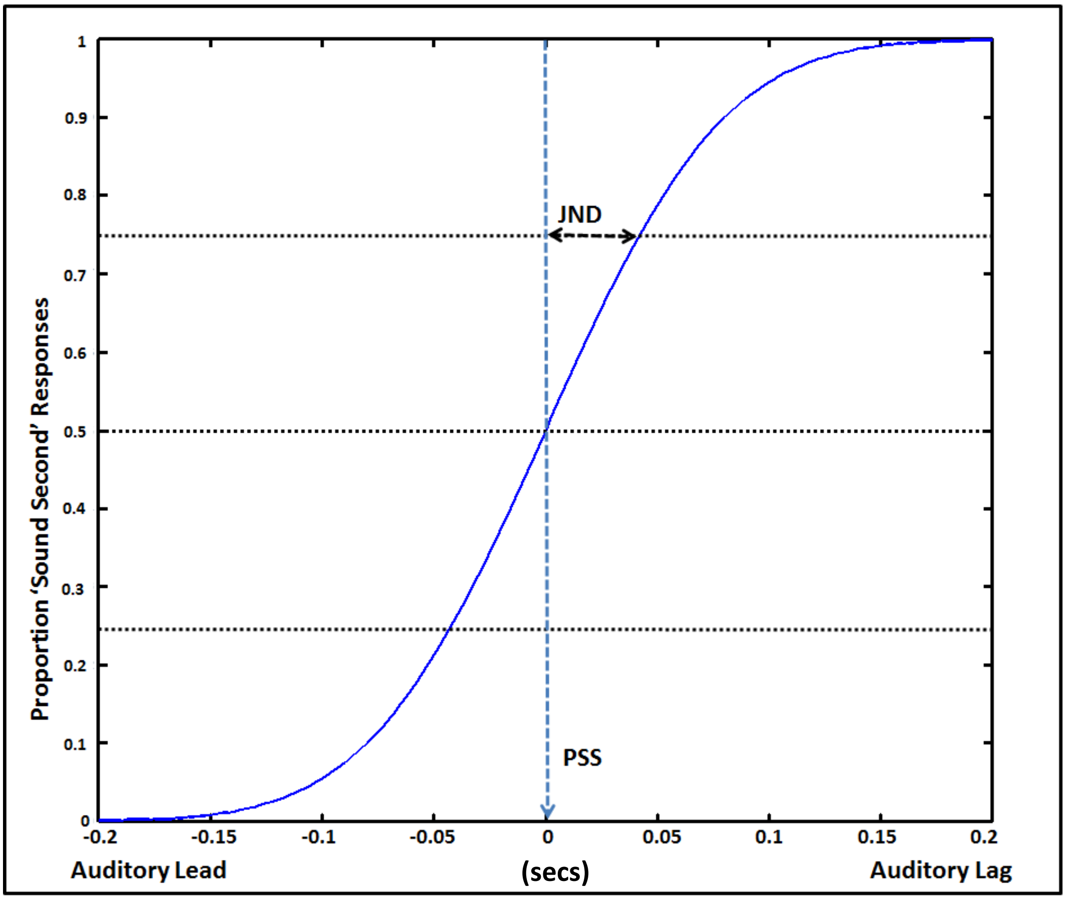


*Figure 1-3: Hypothetical data of a temporal order judgement (TOJ) task. The point of subjective synchrony (PSS) is the asynchrony on the x axis that corresponds to a response proportion of 0.5. Here, it is observed when the auditory and visual information are presented synchronously. The just-noticeable difference (JND) is traditionally estimated by halving the difference between the AV asynchronies at which response rates are 0.75 and 0.25. Here, the JND is roughly 50ms.*

For TOJs, the JND is calculated by halving the distance between the SOAs at which the participant responded 'sound first' 25% and 75% (e.g. Spence et al., 2003) and reflects the slope of the cumulative function (see Figure 1-3 on previous page). The JND is often used as a measure of AV temporal resolution (Marja Laasonen, Service, & Virsu, 2002) which is the ability to discriminate between synchronous and asynchronous AV stimuli (Vatakis & Spence, 2008b). Alternatively, the standard deviation of the function can be used to represent sensitivity to temporal order or synchrony (Yarrow et al., 2011)

#### 1.1.1.1 THE EXPLICIT POINT OF SUBJECTIVE SYNCHRONY (EPSS)

Values of the ePSS derived using SJs tend to vary between -10ms (Fujisaki, Shimojo, Kashino, & Nishida, 2004) and as much as 120ms (Dixon & Spitz, 1980). Negative values indicate that auditory information is leading the visual when synchrony is maximally perceived. Average PSS values derived using TOJs generally tend to be more negative than those derived from SJs (see Figure 1-4 below for comparison), in that the auditory stimulus needs to lead the visual in order for the observer to make chance-level decisions about temporal order (van Eijk, Kohlrausch, Juola, & van De Par, 2008). Measures of the TOJ PSS have been reported to lie anywhere between -84ms (Vatakis & Spence, 2006) to +75ms (Zampini, 2003) for simple stimuli like beeps and flashes. The PSS derived from SJs is on average more likely to be either closer to physical AV synchrony or visually lagging, so that the visual

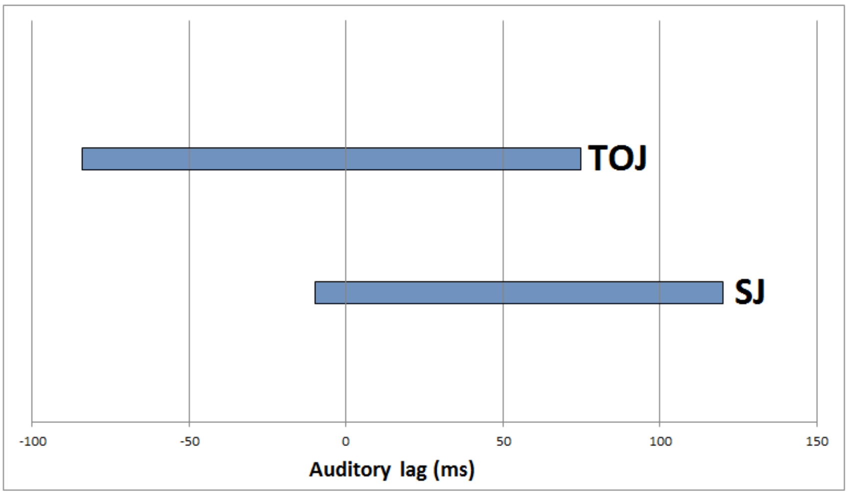


*Figure 1-4: Range of PSS values reported in literature for simple stimuli such as beeps and flashes. Negative Auditory lag indicates that the auditory stimulus was presented first.*

information needs to lead in order for participants to view the two streams as synchronous. As this concurs with what is known regarding the neural transmission of visual and auditory information, SJs have been thought to be the preferred method of measuring the PSS (van Eijk et al., 2008).

PSS values, in particular those derived from TOJs, seem to depend on the modality on which attention is placed (Massimiliano Zampini, Shore, & Spence, 2005), stimulus complexity (Vatakis & Spence) and intensity. Observer stimuli distance (Sugita & Suzuki, 2003) can also lead to different measures of PSS, although this view is debated (see Vroomen & Keetels, 2010 for a review). The PSS is also affected by where the observer sets their decision criteria when making temporal judgements. As a result, two individuals with different measures of PSS may actually process AV information at the same relative timing but may place their decision criteria at different AV asynchronies.

#### 1.1.1.2 DISCRIMINATION OF AUDIOVISUAL TEMPORAL ORDER AND ASYNCHRONY

The ability to discriminate the temporal order of AV events depends on stimulus complexity and on whether the two events originate from the same or from different locations. For example when simple beeps and flashes originate from different locations, the order in which they are presented, i.e. sound first or light first, can be reliably discriminated when the two are separated by as little as 20ms to 58ms but when they are presented in the same place, discrimination becomes poorer and JND measures range between 36ms-90ms (Eskes, Klein, Dove, Coolican, & Shore, 2007; Hirsh & Sherrick, 1961; Keetels & Vroomen, 2005; Spence et al., 2003; Zampini, 2003). The temporal order of more complex stimuli like AV musical notes or monkey vocalisations is reliably judged at slightly longer AV asynchronies, ranging from 58ms to 109ms (Vatakis, Ghazanfar, & Spence, 2008; Vatakis & Spence, 2006, 2008b). AV speech stimuli such as syllables and words has been reported to be reliably judged correctly when the audio and video streams are separated by as little as 59ms to as much as 150ms (Salvador Soto-Faraco & Alsius, 2007; Vatakis, Ghazanfar, et al., 2008; Vatakis & Spence, 2006, 2007, 2008b). These values are illustrated in Figure 1-5, below.

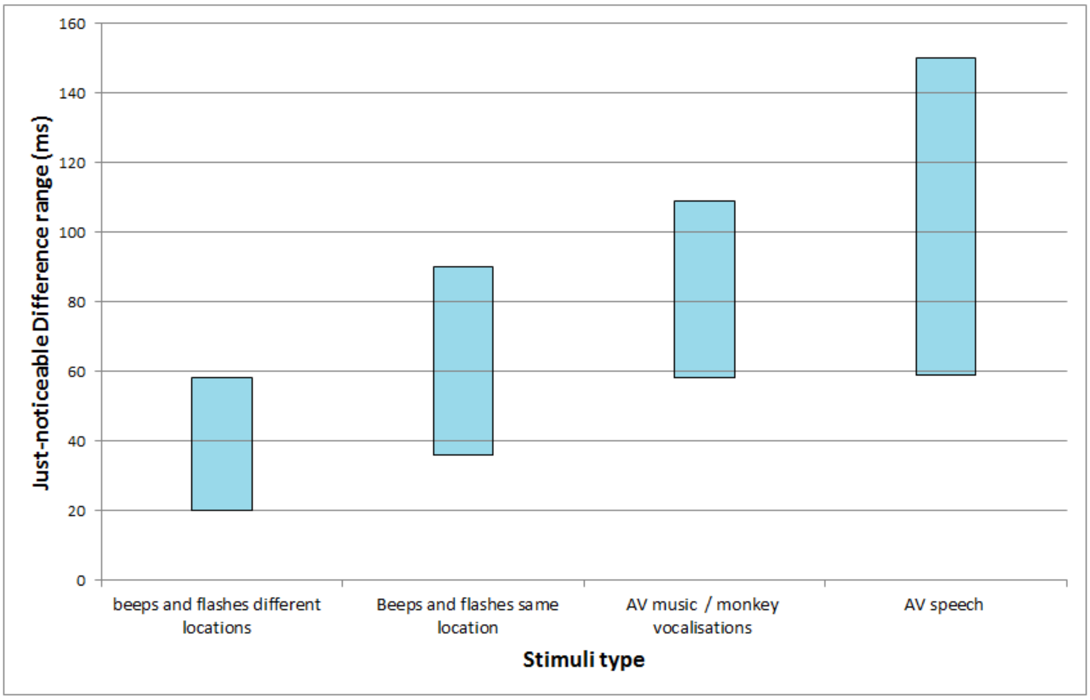


*Figure 1-5: Range of just-noticeable difference values reported in literature, for different types of stimuli.*

The methodology used to derive estimates of the ability to discriminate AV synchrony from asynchrony using SJs varies from study to study. For example, some examine the temporal window of AV synchrony (e.g. Cook, van Valkenburg, & Badcock, 2011; Petrini et al., 2009; van Wassenhove et al., 2007) which represents the range of AV asynchrony within which observers judge auditory and visual events as synchronous. This measure can be computed in different ways. For example, Petrini et al. (2009) and Vroomen and Stekelenburg (2011) estimated the window of synchrony by computing the standard deviation of the normal distribution curve fitted to the raw data. Petrini et al. (2009) reported the width of the window for drumming actions and sounds to range from 100 to 200ms depending on the observers' musical expertise, with experts having smaller windows. Vroomen and Stekelenburg (2011) reported that the window of synchrony was approximately 170ms for both speech and sine-wave speech replicas. van Wassenhove et al. (2007) defined the window of synchrony as the range between the two points at which synchronous responses decreased significantly on either side of the ‘synchronous’ response distribution. These points were defined by an asymmetrical double sigmoid function fitting procedure. The authors reported that congruent AV speech stimuli were perceived to be synchronous when presented at asynchronies anywhere between 73ms auditory-leading to 131ms auditory-lagging asynchronies, totalling a temporal synchrony window width of 204ms (van Wassenhove et al., 2007). Other studies measured the distance between the two points at which stimuli are judged synchronous 75% of the time (Soto-Faraco & Alsius, 2009) and reported that observers judge AV speech as synchronous at asynchronies anywhere between 70ms auditory-leading and 234ms auditory-lagging. Another method of operationalising AV temporal sensitivity is by deriving the slopes of the two separate cumulative functions fitted to each side of the raw SJ data, split by the maximum ‘synchronous’ response proportions (van Eijk et al., 2008). These measures are difficult to compare because although they represent AV temporal sensitivity, they are derived in different ways and are not necessarily analogous to one another.

### 1.1.2 AUDIOVISUAL SYNCHRONISATION DURING AUDIOVISUAL INTEGRATION

Psychophysical research has predominantly employed explicit timing judgements to measure the characteristics of AV synchrony perception and the mechanisms that underlie it. AV synchronisation does however also occur implicitly, as part of AV integration (Martin et al., 2012). Whilst this type of AV synchronisation has been measured more in the context of AV integration and its temporal constraints, it has not been discussed or used much in the context of AV synchrony processing. Implicit synchronisation refers here to a processing of AV timing which does not require conscious effort on the part of the observer. It also reflects the way in which implicit measures analogous to the PSS and JND usually measured using explicit temporal judgements can be derived. The observer is not asked to attend to the temporal characteristics of AV stimuli, nor are they asked to make explicit judgements on AV relative timing. Auditory and visual stimuli are still presented at

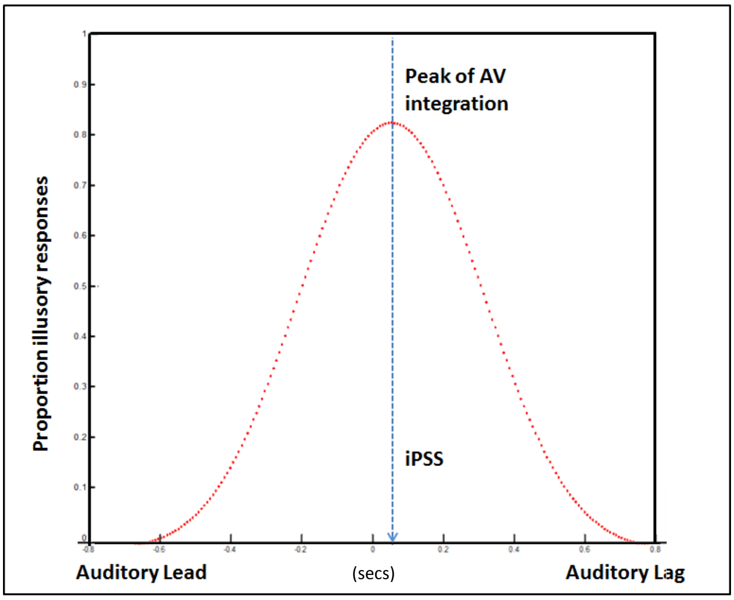


*Figure 1-6: Proportion of illusory responses plotted as a function of auditory lag (negative values indicate auditory lead), and fitted with a psychometric function. The asynchrony at which AV integration is maximal, denoted here iPSS, can then be read from the x axis and taken to represent the point of implicit subjective synchrony.*

various SOAs, but the perceptual judgement required of the observer is related to the ‘what’ rather than to the ‘when’ aspect of the stimuli. Measuring then plotting audiovisual integration as a function of AV asynchrony, instead of temporal judgements, results in a bell shaped curve similar to that which results from SJs (e.g. Asakawa, 2008; Fujisaki, Shimojo, Kashino, & Nishida, 2004). As illustrated in Figure 1-6 above, the asynchrony at which AV integration maximally occurs can then be read from the point on the ‘x’ axis which corresponds with the peak of the function and taken to represent the implicit point of subjective synchrony (henceforth iPSS). Implicit temporal sensitivity can be derived from measures which reflect the degree to which audiovisual integration occurs beyond its optimal asynchrony, such as the width of the curve, or the slopes of its sides. The slopes can be estimated using the standard deviation of the function(s) fitted to the data or using methods in which the JND or window of synchrony is derived from SJ data, described in section 1.1.1.2.

### 1.1.2.1 THE MCGURK ILLUSION

Audiovisual integration can measured using a variety of methods. Probably the most famous AV speech illusion is the McGurk-MacDonald illusion (henceforth the McGurk effect; McGurk & MacDonald, 1976). Introducing a mismatch between the information conveyed by auditory and visual streams can give rise to a subjective

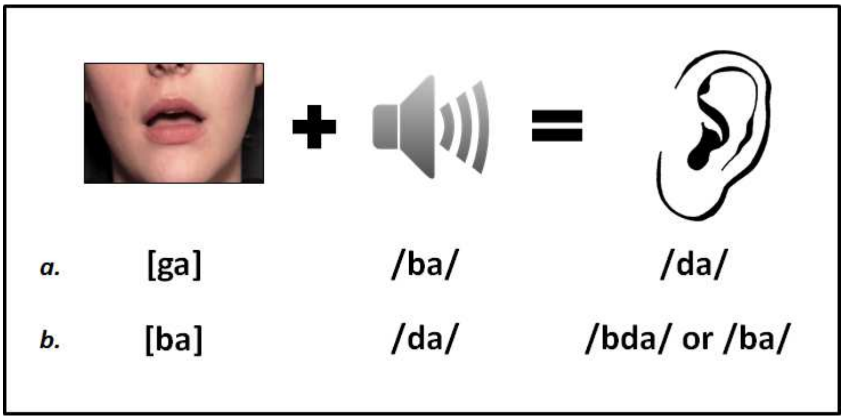


*Figure 1-7: The McGurk Illusion.* ***a.*** *depicts the Fusion illusion in which the AV information is fused into a new auditory percept and* ***b.*** *depicts the combination illusion in which the observer reports hearing a combination of the visual and auditory information, or the phoneme conveyed by the lip-movements only.*

percept which is either qualitatively different than both the visual and auditory components of the stimulus presented, or reflect a combination of the two. For example, as depicted in Figure 1-7 a. above, when presented with the auditory phoneme /ba/ and the incongruent lip movement [ga], observers often report hearing /da/; the same effect is seen with the phonemes /pa/ and [ka], which results in the subjective auditory percept /ta/. For presentations of phonemes such as auditory /ba/ and visual [da], observers will often report hearing /bda/, a combination of the information conveyed by the two modalities, as seen in Figure 1.7 b. On average, this illusion occurs maximally roughly when the auditory and visual streams are presented synchronously and decreases in strength as AV asynchrony increases (Soto-Faraco & Alsius, 2009; Soto-Faraco & Alsius, 2007; van Wassenhove et al., 2007).

The average asynchrony at which the McGurk effect peaks has been reported to be around the point when the auditory stream lags the visual by approximately 60-70ms (e.g. Asakawa, 2008; Munhall, Gribble, Sacco, & Ward, 1996b; van Wassenhove et al., 2007), suggesting that AV synchrony may not be the optimal condition for AV integration. The McGurk effect also does not seem to entirely break down once the auditory and visual streams are separated by seemingly large asynchronies such as 233ms (Asakawa, 2008), 240ms Munhall et al., 1996) and

267ms (van Wassenhove et al., 2007, see Figure 1.8 for illustration of the temporal profie of the McGurk illusion).

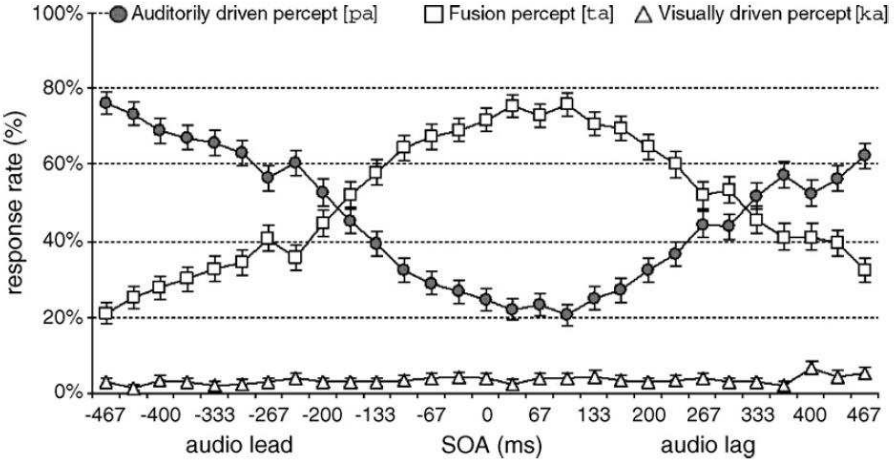


*Figure 1-8: Figure borrowed from van Wassenhove et al. (2007), illustrating the temporal profile of the McGurk illusion. As absolute asynchrony increases, illusory responses decrease and veridical auditory driven responses increase.*

The temporal window of AV integration in the McGurk effect has been reported to be on average around the same width as the window of temporal integration measured using synchrony judgements, which is around 200ms, but has been reported to be smaller than the window of asynchrony for some stimulus combinations (e.g van Wassenhove et al., 2007). Another study (Soto-Faraco & Alsius, 2009) however reported illusory McGurk responses to occur equally as or more often than veridical responses within a much wider temporal window, ranging between 320ms auditory-leading to 480ms auditory-lagging.

### 1.1.2.2 THE STREAM-BOUNCE ILLUSION

Implicit synchronisation can also be measured using AV illusions that do not rely on speech stimuli. For example, in the Stream-Bounce illusion (Sekuler, Sekuler, & Lau, 1997) depicted in Figure 1-9 on the next page, the perceived trajectory of a pair of visual stimuli is affected by the occurrence of a single beep. Two identical disks are displayed at each corner of a computer display. Each disk then begins to move along a downward diagonal trajectory, crossing over the other disk at the midpoint of the display, each of them eventually reaching the opposite corner of the screen.

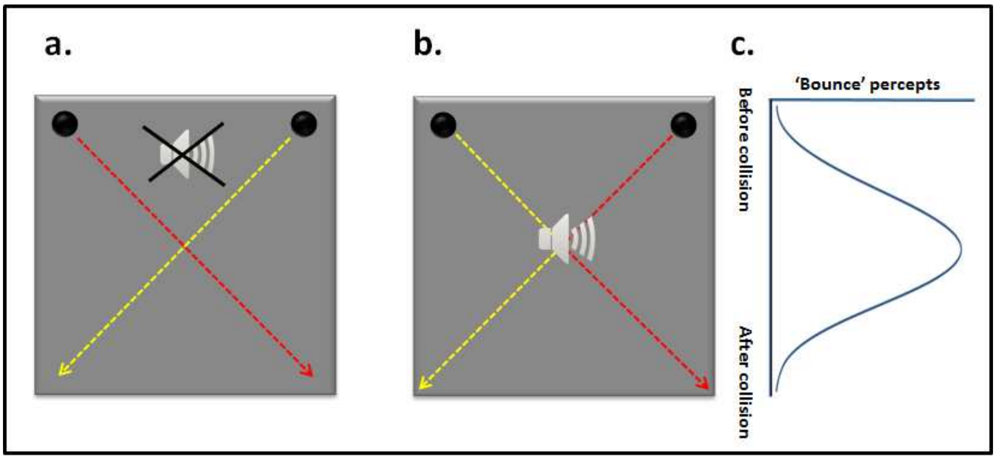


*Figure 1-9: Illustration of the Stream-Bounce illusion. Red and yellow dotted lines represent perceived trajectory **a.** scenario where there is no sound and a corresponding percept of streaming. **b.** scenario where the sound occurs at the same time as the point at which the disks meet in the centre of the display and the corresponding percept of bouncing. **c.** Hypothetical probability of 'Bounce' percept as a function where the disks are located on the display when the beep occurs.*

The disks' actual trajectory is towards the opposite side of the screen, as the two disks become one in the centre of the display and therefore 'stream' through one another. Despite this, the display can also be interpreted as the disks bouncing against each other in the center, resulting in a perceived change of trajectory towards the same side of the screen. The addition of a single auditory beep when the two disks are positioned at or around the midpoint, increases the likelihood of a 'bounce' percept. As the asynchrony between the beep and the collision point of the disks increases, the proportion of bounce responses decreases (Fujisaki et al., 2004), which is qualitatively similar to the temporal profile of the McGurk illusion.

### 1.1.3 IMPLICIT VERSUS EXPLICIT TEMPORAL JUDGEMENTS

Although implicit and explicit synchrony judgements produce similar data that can be fitted in the same way and compared, this does not necessarily mean that the two types of tasks measure analogous temporal processing ability. As Chapter 2 and 3 will discuss and demonstrate, implicit and explicit temporal processing may not rely on common underlying mechanisms (Soto-Faraco & Alsius, 2009; Soto-Faraco & Alsius, 2007). This is problematic for interpreting research which uses explicit temporal judgement paradigms to operationalise AV integration (Petrini et al., 2009; Vatakis & Spence; 2007) or which discusses findings related to the perception of synchrony to AV integration (Vatakis & Spence, 2007). For instance, Vatakis and

Spence (2007) investigated the effects of the unity assumption on the temporal window of AV integration of speech stimuli, and operationalised the latter using measures of the JND from temporal order judgements. Navarra et al. (2005) also measured performance on temporal order judgements after exposure to asynchronous speech and took poorer discriminability of AV temporal order to represent a widening of the temporal window of audiovisual integration.

The windows of AV synchrony and AV integration seem to be different, suggesting that the mechanisms underlying the two tasks might also differ. Soto-Faraco & Alsius (2007) measured temporal order perception as well as AV integration using the McGurk effect. The window between auditory-leading and lagging asynchronies within which responses were correct less than 75% of the time was taken to represent the range of AV asynchronies within which participants were uncertain of AV temporal order, and therefore perceived the stimuli to be synchronous. This range was reported to be between 94ms auditory-leading and 208ms auditory-lagging asynchronies. On the other hand, the window within which McGurk responses were equal to or greater than veridical responses (i.e., the window of AV integration) was reported to lie between 400ms auditory-leading and 480ms auditory-lagging asynchronies. Therefore, the window of synchrony for TOJs was 304ms whereas the temporal window of McGurk integration was 880ms, which is more than twice as large as the window of AV synchrony.

Petrini et al. (2009) investigated the effects of expertise on synchrony perception but refer to the window of synchrony as the window of AV integration. The authors also discuss their rationale and findings interchangeably in the context of audiovisual integration and synchrony perception. Using explicit temporal judgements to represent AV integration would be acceptable if explicit temporal judgements and AV integration were reliant on the same underlying mechanisms. However, if these tasks do not reflect the same underlying processes doing so could lead to invalid accounts of AV integration and timing.

AV integration tasks and explicit temporal judgements are also subject to different response biases. AV integration tasks are in theory less prone to decision bias, because the responses given by participants in implicit AV timing tasks do not relate to the timing of the stimuli. For example in SJs, a widening of the window of synchrony could reflect a bias towards responding ‘synchronous’ more often as opposed to a tendency to actually perceive AV stimuli as synchronous when they are asynchronous. This could result from the observer relaxing their decision criteria regarding what is synchronous and asynchronous in order to increase their confidence in their judgements. In contrast, in the McGurk effect there is no intuitive reason as to why participants would use or loosen such criteria within which they would report that they heard the syllable [da] as opposed to [ba] as the task is not directly related to judging AV relative timing. As a result, the temporal window of AV integration may be less ambiguous to interpret than the temporal window of perceived synchrony. Critically, in regards to the interpretation of effects, certain manipulations might result in a widening or shortening of SJ windows of synchrony as a result of shifts in decision criteria, but they might not have the same effect on the temporal window of integration. Therefore, using synchrony judgements to operationalise AV integration might lead to invalid interpretation of the temporal constraints and characteristics of AV integration, and this is especially true if the two tasks measure the performance of two different mechanisms.

Research which investigates the way in which the nervous system might minimize external audiovisual delays by adapting to them is predominantly carried out using explicit temporal order and synchrony judgements, which in principle measure the ability to *explicitly* access and judge AV relative timing (Heron, Hanson, & Whitaker, 2007; Roseboom & Arnold, 2011; Vatakis, Navarra, Soto-Faraco, & Spence, 2008; Vroomen, Keetels, de Gelder, & Bertelson, 2004). For example, temporal recalibration occurs when exposure to asynchronous AV information results in AV streams which were perceived as being asynchronous prior to exposure, to be perceived as synchronous post exposure (Fujisaki et al., 2004; Hanson, Heron, & Whitaker, 2008; Vatakis, Navarra, et al., 2008). This phenomenon is thought to reflect mechanisms which maintain temporal coherence across modalities, in spite of internal and external factors which create temporal disparities between sensory information that occur synchronously. Such factors include differences between transduction and propagation of auditory and visual signals, or in the time taken by auditory and visual information to reach an observer (Fujisaki et al., 2004; Hanson et al., 2008; Keetels & Vroomen, 2007). These temporal adaptation effects are often tested using SJs (Roseboom & Arnold, 2011) and TOJ tasks (Luca, Machulla, & Ernst, 2009; Vatakis, Navarra, Soto-Faraco, & Spence, 2007) and less often using AV integration paradigms (but see Asakawa, 2008 and Fujisaki et al., 2004). Although some of the studies restrict the discussion of the results to mechanisms underlying explicit synchrony perception (Luca et al., 2009; Roseboom, Nishida, & Arnold, 2009), others (e.g. Vatakis et al., 2007) generalise theoretical conclusions to temporal mechanisms underlying AV integration, which is problematic because it has not yet been established whether AV integration and explicit temporal judgements rely on the same temporal mechanisms.

To summarise, the foregoing studies used explicit timing judgements as means of measuring AV integration or have generalised findings obtained using explicit temporal judgements to temporal mechanisms underlying AV integration. This would not be problematic if explicit and implicit synchronisation processes are indeed supported by the same underlying mechanisms. However, whether AV integration and explicit AV temporal judgements are supported by common or distinct mechanisms has not yet been directly tested until now. This issue is discussed is more depth in Chapters 2 and 3. In Chapter 2, the question is addressed by examining whether measures of subjective synchrony derived from implicit and explicit AV temporal judgements are consistent within individuals. In Chapter 3, the question is addressed by examining whether the ability to discriminate between synchronous and asynchronous AV information across implicit and explicit temporal judgements is statistically dependent or independent and whether it is related to variability in the structure of similar or distinct brain clusters. These chapters use a different approach to the one that has been adopted by literature so far. Instead of examining data across participants, this thesis examines whether and how individual differences in performance across these qualitatively different tasks covary within individuals. The rationale and benefits of this approach is discussed in the next section.

## 1.2 INDIVIDUAL DIFFERENCES

A relatively large proportion of what is known about the temporal profiles of AV integration and the perception of AV synchrony and temporal order is based on data which have been averaged out across participants. However, the temporal profiles of AV integration and of AV synchrony perception (Martin et al., 2012), measures of the implicit (Freeman et al., 2013) and explicit PSS (Boenke, Deliano, & Ohl, 2009; Stone et al., 2001) and susceptibility to AV illusions (Nath & Beauchamp, 2012) have been reported to be subject to individual differences. For example, studies which focus on AV integration have often had to exclude individuals because they were not susceptible to illusions such as the McGurk effect (e.g. Nath & Beauchamp, 2012; Skipper, van Wassenhove, Nusbaum, & Small, 2007) and some have even found that brain activity in multisensory areas differs across susceptible and non-susceptible individuals (Szycik, Stadler, Tempelmann, & Münte, 2012).

One study which examined individual differences in the perception of synchrony was carried out by Boenke et al., (2009). The authors looked at whether individual differences had an effect on the degree to which stimulus duration shifted the PSS of AV stimuli, after finding that an Analysis of Variance revealed no overall effect of stimulus duration on the PSS, despite some individual data points showing large shifts. Correlational analyses were carried out on measures of the PSS and the degree to which individual estimates of these measures shifted as a result of stimulus intensity and duration. This study firstly demonstrated that the PSS is subject to wide inter-individual variability and can range from anywhere between around 120ms auditory-leading and 150ms auditory-lagging AV asynchronies. In addition, the authors measured the size of the effect of stimulus duration on the PSS in each participant and found that the shifts in PSS also varied from 75ms towards auditory-leading and 75ms towards auditory-lagging. More importantly correlational analyses revealed that the size and direction of this effect in a given individual depended on size of their PSS and its direction. Individuals with a large auditory-leading PSS showed a large shift away from AV synchrony and towards even larger auditory-leading asynchrony whereas individuals with a large auditory-lagging PSS showed a large shift away from zero but in the opposite direction, towards larger auditory-lagging asynchrony. For individuals with a veridical PSS, the shift in PSS was minimal.

The study carried out by Boenke et al. (2009) demonstrates not only that there is wide variability in the PSS and in the degree to which it affected by stimulus characteristics, but that the two are not orthogonal. Individual variability in the PSS is related to the degree to which factors such as stimulus duration modulate it. Thus, an ‘individual differences’ approach using correlational analysis revealed an effect of stimulus duration on the PSS which otherwise would have not been revealed using an analysis of group averages because the effect itself was modulated by individual differences in subjective synchrony. To summarise, examining individual variability can reveal subtle characteristics of AV temporal processing which might otherwise be discarded and therefore concealed by group averaging analyses.

Another study which demonstrates the benefits of using an individual difference approach was carried out by van Eijk, Kohlrausch, Juola, & van De Par (2010). The authors examined the relationship between individual differences in the TOJ PSS and sensitivity in discriminating between synchronous and asynchronous AV information during SJs, to test the hypothesis that the PSS obtained in a TOJ task is shifted towards the AV asynchrony to which individuals are most sensitive when judging synchrony. The authors fitted cumulative Gaussians to each side of the synchrony judgement distribution for each participant and computed slopes for each curve, which represent temporal sensitivity. A ratio was then computed which represented whether participants were more sensitive to asynchrony when the auditory led the visual information or whether they were more sensitive when the auditory lagged. They found that the PSS was related to this ratio, meaning that participants with a higher sensitivity for audio-leading asynchrony also tended to have audio-leading PSS values, whereas participants with higher sensitivity for audio-lagging asynchrony tended to have more audio-lagging values of the PSS. This relationship between sensitivity during SJs and the PSS derived from TOJs would not have been as easily revealed using group averages.

An individual differences approach might therefore reveal subtle differences or relationships between implicit and explicit temporal processes which have not yet been revealed by studies employing a group averaging approach. For example, Chapter 2 of this thesis examines whether individual differences in the PSS are consistent within individuals across implicit and explicit temporal judgements, or whether they are statistically independent from one another. Part of Chapter 3 carries out similar analyses on the ability to discriminate between synchronous and asynchronous AV information across implicit and explicit temporal judgements. Positive correlations between specific aspects of AV temporal processing performance across implicit and explicit judgements would indicate that these different types of judgements are supported by common underlying mechanisms. In contrast, if these measures are statistically independent, it could be an indication that these different types of judgements might be supported by distinct underlying mechanisms.

Why individuals might differ from one another in their ability to synchronise or to integrate AV information in the first place is unclear. Individual differences in brain function have been shown to reflect individual variability in perception and behaviour, but it is unclear from functional correlates whether it is activity that leads to perception and/or behaviour, or vice versa. Nath and Beauchamp (2012) found that higher susceptibility to the McGurk illusion, measured outside of the scanner correlates with higher BOLD signals in the Superior Temporal Sulcus (STS) during exposure to incongruent McGurk stimuli. The STS is known to be one of the critical underlying structures for AV integration as disruption of this area using transcranial magnetic stimulation (TMS) significantly decreases the McGurk effect (Beauchamp, Nath, & Pasalar, 2010). The findings demonstrate that individual differences in AV integration are reflected by individual variability in the strength of

BOLD activation in multisensory cortex. However but the causal direction of the relationship between BOLD signal and perception is unclear without the use of additional methods such as TMS.

Hipp, Engel, & Siegel (2011) also demonstrated that individual differences in AV perception were reflected in individual variability in brain activity, but with the use of EEG rather than fMRI. The authors examined the role of long-range gamma-band oscillatory synchronisation within what they labelled a ‘centro-temporal’ network in AV integration, measured by the Stream-Bounce illusion (Sekuler et al., 1997). Individual susceptibility to the illusion was correlated to the degree to which gamma-band synchronisation increased across trials in which the auditory stimulus was integrated with the visual, relative to trials in which auditory information had no effect on visual perception. Less susceptible individuals tended to have larger differences in gamma-band synchronisation across the two percepts whereas more susceptible individuals tended to have comparable levels of gamma synchronisation across the percepts. This relationship was driven by synchronisation observed during trials in which AV information was integrated, meaning that in general, the less susceptible individuals were to the Stream-Bounce Illusion, the more oscillatory coherence they showed during illusory trials and the more susceptible individual were, the less synchronisation they showed. These results therefore suggest that gamma-band synchronisation within the centro-temporal network is unlikely to be the mechanism that gives rise to the Stream-Bounce illusion because highly susceptible individuals were less likely to show high levels of it. The results suggest that it may however act as a compensatory mechanism acting to facilitate illusory perception in individuals who generally show low susceptibility. This result might not have been revealed if differences in the susceptibility to the Stream-Bounce illusion and in the strength of gamma-band synchronisation across individuals were discarded by only using analysis in which data are averaged, emphasising the need to also take individual differences into account when interpreting data.

The causal relationship between individual differences in perception and individual variability in brain activity is ambiguous. For example, in Nath and Beauchamp’s study, it is not entirely straightforward that increased BOLD signal might lead to a given percept; it is just as likely that perceiving might lead to increased BOLD signal. Greater cortical synchronisation in Hipp et al.'s (2011) study might also be a consequence, rather than a cause of AV integration. Furthermore, even if there was no doubt about the possibility that increased BOLD signals or neural synchronisation were causal factors in perception or behaviour, why some individuals might show greater activity or neural synchronisation would still remain unclear.

## 1.3 VOXEL-BASED MORPHOMETRY

Brain structure is one factor that might account for individual variability in brain activity as well as behavioural variability in AV integration and temporal processing, but this possibility has not yet been addressed in AV research. Brain morphology is particularly interesting to examine as it provides a potentially causal explanation for individual differences in AV integration and temporal processing that is less ambiguous than that which is provided by purely functional correlates. For example, increased grey matter density indicates that a given area is likely to contain more neurons and/or nerve fibres, which might be the resources necessary for less noisy and more efficient computations, as these transmit information through the central nervous system. If better performance correlates with increased grey matter volume, it might explain why some individuals are better than others at integration and synchronising AV information. Such an interpretation is consistent with population coding models which propose that stimuli are represented by the distributions of responses of different neuronal populations, each tuned to different types of information (Averbeck, Latham, & Pouget, 2006). A distribution of responses to AV asynchrony for example might be less noisy if it were produced by a larger population of neurons, thus individuals with larger neuronal populations in areas responsible for AV timing might be better at synchronising AV information than individuals who possess smaller neuronal populations.

Voxel based morphology is a voxel-wise analysis of the local density of grey matter in the brain and has been used predominantly in past investigations to address whether clinical populations differ in terms of brain structure from non-clinical populations (e.g. Boddaert et al., 2004; Chung, Dalton, Alexander, & Davidson, 2004; Dole, Meunier, & Hoen, 2013; Valente et al., 2005). More recently however, VBM has been used in conjunction with an ‘individual differences’ approach in order to examine whether individual differences in behaviour across participants can predict local grey and white matter density, in order to elucidate the neural mechanisms underlying behaviour.

Brain structure variability has been linked to individual differences in visual perception (Kanai, Bahrami, & Rees, 2010), performance in attention (Westlye, Grydeland, Walhovd, & Fjell, 2011) and action selection tasks (van Gaal, Scholte, Lamme, Fahrenfort and Ridderinkhof, 2011), as well in variability in personality traits (DeYoung et al., 2010) and social cognition (Bickart, Wright, Dautoff, Dickerson, & Barrett, 2011).

Given that individual differences in behaviour, perception and traits can be reflected in brain structure variability as demonstrated by the aforementioned studies, Chapter 3 makes use of VBM to investigate whether performance in implicit and explicit AV temporal judgements can be dissociated at the neural level. For example, if individual differences in these tasks correlate with structural variability in common areas it would be an indication that implicit and explicit AV synchronisation might be supported by common neural mechanisms. Alternatively, if the analysis reveals that performance in implicit tasks correlates with structural variability in distinct areas from those related to explicit tasks, it would suggest that they might be supported by different neural mechanisms.

## 1.4 AUDIOVISUAL INTEGRATION AND SYNCHRONISATION, READING ABILITY AND DYSLEXIA

Individual differences in explicit and implicit temporal processing and AV integration might also inform on other, higher cognitive processes which might depend on these basic perceptual mechanisms. Chapter 4 explores whether AV integration and timing are related to reading ability and dyslexia. The chapter examines whether on average, performance in AV temporal processing and AV integration tasks differs between individuals diagnosed with a reading disability and typical readers, as well as whether AV timing and integration is related to reading ability, over and above dyslexia.

The successful formation of AV correspondences between auditory and visual speech is believed to contribute to language development (Teinonen, Aslin, Alku, & Csibra, 2008) and may also be necessary for learning grapheme-phoneme correspondences (Blomert & Froyen, 2010). AV relative timing affects AV integration in children less than a year old, thus even at an early, pre-linguistic stage of development, synchronisation of AV information is important in audiovisual integration (Gogate & Bahrick, 1998; Hollich, Newman, & Jusczyk, 2005). Deficits in temporal processing could affect the ability to form AV correspondences which later may be useful in learning to pair graphemes and phonemes automatically, and may lead to poor reading skills later in life. For example, dyslexia is a neurobiological condition characterised by reading problems such as recognition, spelling and decoding of words, despite otherwise typical reading instruction and educational or professional attainment (Lyon, Shaywitz, & Shaywitz, 2003). Individuals diagnosed with this condition show reduced automaticity in grapheme-phoneme association which is likely to be a result of poor learning of grapheme-phoneme correspondence (Ramus, 2001). This could be partly the result of deficits in AV temporal processing and/or integration, but only a limited amount of research has explored this possibility and the role of audiovisual processing in reading is not well understood. The available research on reading ability, dyslexia and AV processes are briefly summarised here and reviewed in depth in Chapter 4.

Conclusions on whether AV integration is impaired in dyslexia are mixed, possibly due to the small number of studies on the topic. Two behavioural studies have reported that the McGurk effect is intact in dyslexic individuals and concluded that AV integration is unimpaired in the condition (Bastien-Toniazzo, Stroumza, & Cavé, 2010; Campbell, Whittingham, Frith, Massaro, Cohen, et al., 1997), whereas another has found that dyslexic individuals gain less information from visual information during speech-in-noise detection (Ramirez & Mann, 2005), suggesting that AV integration in dyslexia is impaired. Both the McGurk effect and AV speech-in-noise detection are ways in which AV integration is measured, thus these studies suggest opposing conclusions regarding whether AV integration in dyslexia is affected. These studies however presented AV information synchronously and there is a possibility that temporal processing of AV information rather than its integration is affected in the disorder. Hairston, Burdette, Flowers, Wood, & Wallace (2005) found evidence to suggest this; they reported that dyslexic individuals integrate simple beeps and flashes at larger AV asynchronies compared to typical readers, indicating that dyslexic individuals might have different temporal profiles of AV integration compared to typical readers.

Brain imaging studies seem to suggest that at least at the neural level, AV integration processes are impaired in dyslexia. Froyen, Willems, & Blomert,(2011) reported that electrophysiological markers of automatic AV integration of sounds and letters normally observed in typical readers (Froyen, van Atteveldt, Bonte, & Blomert, 2008) are absent in dyslexic children. Widmann, Schröger, Tervaniemi, Pakarinena and Kujala (2012) reported similar findings. This particular study also found impairments at the behavioural level. Dyslexic children showed poorer ability in performing congruence judgements on sound patterns and visual symbols compared to typical readers, and their ERP responses during exposure to incongruent auditory and visual stimuli were different to those of controls. Activity of areas typically classed as multisensory such as the Superior Temporal cortex

(Nath & Beauchamp, 2012; Noesselt et al., 2012) differs in dyslexic children (Blau et al., 2010) and adults (Blau, van Atteveldt, Ekkebus, Goebel, & Blomert, 2009) compared with activity observed in typical readers, when measured during perception of speech sounds and letters which behaviourally does not differ across the groups.

Audiovisual temporal processing in dyslexia and its relationship to reading is less well understood than audiovisual integration in dyslexia. Explicit crossmodal temporal processing of audiotactile and visuotactile stimuli has been reported to be poorer in dyslexia and to correlate with phonological awareness in this group (Laasonen et al., 2002). Explicit processing of crossmodal relative timing has also been reported to deteriorate more with age in dyslexia compared to in typical readers (Virsu, Lahti-Nuuttila, & Laasonen, 2003). The literature on crossmodal temporal processing in dyslexia is however restricted to these two studies, and so far no impairments related to explicit temporal processing of AV information have been reported in the literature. Chapter 4 addresses this gap in literature.

The small amount of research that has investigated AV temporal processing in dyslexia has measured AV temporal processing skills in isolation from AV integration skills, which is problematic because, as discussed in this introduction, AV temporal processing is important for AV integration. Explicit AV temporal processing has only been investigated and reported to be intact by two studies (Marja Laasonen et al., 2002; Virsu, Lahti-Nuuttila, & Laasonen, 2003) using stimuli comprising brief flashes and beeps, and has not yet been investigated using speech stimuli. The majority of existing EEG research into AV processes in dyslexia has used children, making it unclear whether differences found in brain activity are persistent into adulthood or whether development of AV processing is simply delayed in dyslexia.

Using a combination of group averaging complemented by analyses of individual differences, Chapter 4 examines whether there are differences between dyslexic and typical readers in speech and non-speech AV integration and temporal processing skills, as well as how these skills relate to specific aspect of reading ability across the entire sample. The chapter thus combines an ‘individual differences’ approach with a group averaging one in order to identify whether performance in AV temporal processing can be linked to reading ability, based on differences between typical readers and those with a formal diagnosis of dyslexia, as well as on the relationships between performance in specific aspects of reading impairment and different aspects of AV processing.

## 1.5 SUMMARY AND THESIS OUTLINE

To summarise, audiovisual synchronisation is likely to be a difficult task as the nervous system is faced with various delays produced by external and internal factors, in the relative arrival time of auditory and visual information which at source occur synchronously. Despite this we seem to, on average, integrate and time AV information correctly. A distinction can be made between AV synchronisation tasks which require that attention is directed towards AV relative timing and tasks which do not. These are referred to as explicit and implicit AV synchronisation is the current thesis. Explicit temporal perception can be measured as a function of AV asynchrony using explicit timing judgements such as TOJs and SJ, whereas implicit synchronisation can be measured in the same way using AV illusions such as the McGurk effect and the Stream-bounce illusion. Subjective synchrony and AV integration are subject to wide individual differences, but despite this, much of what is known about the temporal profiles of AV integration and the perception of AV synchrony is based on group averages. Examining individual variability in these processes can reveal subtle characteristics of these processes which might otherwise be discarded by group averages. Chapter 2 investigates whether audiovisual synchronisation and integration are based on common or distinct timing mechanisms by examining whether individual differences in estimates of subjective synchrony derived from explicit timing tasks covary with individual variability in the AV asynchrony optimal for AV integration.

Individual differences in brain activity reflect individual variation in performance of AV integration tasks and might even underlie this variability. However, what determines differences in brain activity remains unclear. Differences in brain structure might also underlie behavioural variability and provide a less ambiguous brain-behavioural relationship compared to purely functional correlates. Whilst structural variability has been shown to reflect individual variation in other behavioural measures, such as those which quantify visual perception and executive function, personality traits and social cognition, no research has yet investigated the structural correlates of AV integration and temporal processing. Chapter 3 assesses whether performance in AV temporal discrimination across implicit and explicit tasks is related to individual differences in the structure of the same or distinct anatomical brain areas.

Successfully forming AV correspondences may also be necessary for learning grapheme-phoneme correspondences during reading development, but little is known about the relationship between reading ability and individual differences in AV integration and temporal processing and about their potential contribution to reading impairments in dyslexia and in typical readers. Chapter 4 compares performance in AV temporal processing and AV integration across dyslexic and typical readers, and correlates individual differences in this performance with different aspects of reading ability across all readers and examines whether reading ability is related to AV temporal processing and integration, over and above dyslexia.

# 2 CHAPTER 2: DISUNITY ACROSS IMPLICIT AND EXPLICIT ESTIMATES OF SUBJECTIVE SYNCHRONY

## 2.1 INTRODUCTION

The current chapter investigates whether audiovisual (AV) synchronisation during explicit temporal judgements and AV integration are based on common or distinct timing mechanisms. The chapter addresses whether AV information needs to be explicitly perceived as synchronous in order to be integrated, by examining the relationship between estimates of subjective synchrony derived from explicit timing tasks (henceforth the *explicit* point of subjective synchrony: ePSS) and the AV asynchrony optimal for AV integration (henceforth the *implicit* point of subjective synchrony: iPSS). Across two experiments employing a dual-task paradigm, AV integration is measured as a function of AV asynchrony concurrently with subjective AV synchrony. In Experiment 2.1, AV integration is measured using the McGurk illusion (McGurk & MacDonald, 1976), whereas in Experiment 2.3 it is measured using the Stream-bounce illusion (Sekuler et al., 1997). Explicit subjective synchrony is measured separately using both temporal order judgements (TOJ) and synchrony judgements (SJ) across different conditions.

To briefly recap, the use of *explicit AV synchronisation* in this thesis refers to the process underlying conscious awareness of relative AV timing, during explicit judgements of AV synchrony or temporal order. The term *implicit AV synchronisation* refers to a process which does not require conscious awareness of AV synchrony or asynchrony on the part of the observer. The terms also reflect the different ways in which estimates of subjective synchrony are measured. To measure estimates of the explicit PSS (ePSS), observers are asked to make explicit judgements on the synchrony or temporal order of AV stimuli. When measuring the implicit PSS (iPSS), observers are not asked to attend to, nor are they asked to judge the temporal properties of AV stimuli. Instead, the judgement made is based on the observer's perceptual experience of the auditory or visual stimuli when exposed to an AV illusion, at various AV asynchronies. The iPSS is then represented as the asynchrony at which the maximum proportion of illusory responses, and therefore AV integration, is observed (e.g. Asakawa, 2008; Fujisaki et al., 2004; van Wassenhove et al., 2007).

### 2.1.1 THE CASE FOR COMMON TEMPORAL MECHANISMS ACROSS DIFFERENT AUDIOVISUAL PROCESSES

A widely accepted view in AV research is that AV integration is contingent upon the observer's assumption that visual and auditory information belongs together. This is known as 'The unity assumption' (Welch & Warren, 1980) and can be promoted by low-level stimulus characteristics such as spatio-temporal coincidence, through both top-down and bottom-up processes (Vatakis, Ghazanfar, et al., 2008; Vroomen & Keetels, 2010a; Welch & Warren, 1980). Temporal coincidence is said to be a good indication that two events have been caused by a common source and that they belong together (Einhorn & Hogarth, 1986). Thus, under this theoretical position (illustrated in Figure 2.1, below), perceiving that an auditory and a visual event occurred at the same time can lead to the assumption that the stimuli belong together, and the two will be subsequently integrated into a unitary AV percept.

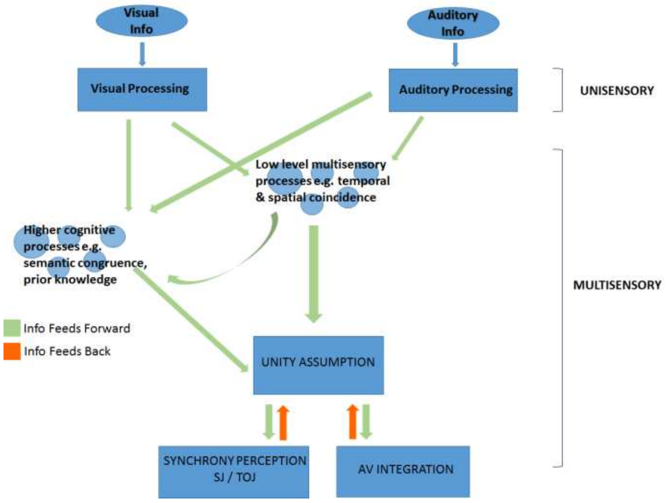


*Figure 2-1: Illustration of the unity assumption.*

The relationship between the perception of temporal coincidence and the unity assumption is said to be bidirectional. In other words, whilst temporal coincidence can encourage observers to make an assumption of unity in certain conditions, the unity assumption can in turn affect the perception of synchrony (Spence, 2007; Vatakis, Ghazanfar, et al., 2008). This is known as the ‘unity effect’, in which AV information that has been integrated is said to be assumed to belong together and therefore perceived as synchronous. This view is motivated by findings which show that asynchronous AV information is more likely to be perceived as synchronous when the visual and auditory streams are congruent and therefore likely to have been integrated, compared to when they are incongruent and thus unlikely to have been integrated (Vatakis & Spence, 2007). Vatakis and Spence created incongruent AV stimuli by either switching gender in one modality so that a male face was presented with a female voice and vice versa, or by presenting the visual stream of an uttered word with the auditory stream of another word. The authors argued that because congruent AV stimuli were more likely to be integrated, they were also more likely to promote the assumption that they belonged together. Conversely, incongruent AV speech stimuli were unlikely to be integrated and thus would not promote the assumption of unity. The Just Noticeable Difference (JND) was measured for both congruent and incongruent AV stimuli. This measure was found to be on average larger for congruent, compared to incongruent stimuli, suggesting that participants found it harder to judge the temporal order of AV events when they promoted an assumption of unity compared to when they were incongruent and did not.

Altogether, this evidence suggests that we benefit from, and strive to achieve unity. Based on this, it would be predicted that the asynchrony at which audiovisual information is perceived to be synchronous will correlate positively with the asynchrony optimal for AV integration.

However, a further study carried out by Vatakis and Spence (2008a) using non-speech, music stimuli did not replicate the unity effect. Neither did a study using the same paradigm with non-speech monkey vocalisations (Vatakis, Ghazanfar, et al., 2008). On the basis of this series of studies (Vatakis, Ghazanfar, et al., 2008;

Vatakis & Spence, 2007, 2008a), Vatakis and Spence (2008) argued that the unity assumption facilitates AV temporal integration of speech signals only. The authors argued that this could possibly owe to a differential distribution of top-town and bottom-up processing through which the unity assumption is promoted across speech and non-speech processing. The authors also argued that speech processing is likely to lead to a 'special' mode of perception, more likely to promote unity compared to non-speech modes of perception (Vatakis, Ghazanfar, et al., 2008). According to this line of reasoning, positive relationships between implicit and explicit subjective synchrony would be stronger for AV speech stimuli.

Under assumptions of unity, temporal coincidence is an indicator of common cause, which increases the likelihood of AV integration. AV integration in turn affects whether two events are attributed to a common cause, influencing whether the two are perceived as synchronous. There is not however clear and direct empirical evidence for the link made between perceiving synchrony and assuming that two events belong together. Research has shown that as the delay between a visual and an auditory stimulus increases, the decision that the two have been caused by the same event decreases, and that a similar pattern is also observed with judgements of AV synchrony (Guski & Troje, 2003; Lewald & Guski, 2003). Correlational analysis has not yet been carried out between the (a)synchronies at which AV synchrony is most likely to be perceived and the (a)synchrony at which a common cause is most likely to be attributed to AV stimuli. Such analyses would be necessary to determine whether individuals are most likely to perceive synchrony between auditory and visual information when they are also most likely to perceive the stimuli as belonging together. Across different experiments, but using the same participants, Lewald & Guski (2003) measured perceived phenomenal causality and subjective synchrony of simple light and sound pulses, as a function of AV asynchrony. In one task, participants had to judge the likelihood that the auditory and visual stimuli had a common cause and in another they judged the degree to which the two were synchronous. In both tasks participants used a scale from 1 to 9 to indicate their subjective perception of the stimuli. The averaged ratings plotted as a function of task and asynchrony, as shown in Figure 2.2. The point of subjective synchrony (PSS) and the asynchrony which yielded the highest average rating of common cause were on average similar; they both occurred when the visual stimulus was presented slightly earlier than the auditory stimuli. However, estimates of these asynchronies were not statistically compared and no correlational analysis was carried out across the two measures within participants, so it is unclear whether the stimuli had to be perceived as synchronous in order to be attributed a common cause. Lewald and Guski qualitatively compared the temporal profiles of the two types of judgements and noted that these were not identical. Participants were likely to judge two stimuli as having a common cause at AV asynchronies they could reliably detect as being asynchronous. Note that the opposite pattern of results has been reported for the profiles of AV integration in the McGurk effect and of perceived AV speech synchrony. van Wassenhove et al.'s (2007) results (Figure 2.3) for one of the stimulus combinations used in their study to measure AV integration and synchrony perception showed that the profile of AV synchrony was on average wider than the profile of AV integration, which suggests that participants did not integrate all AV pairs that they

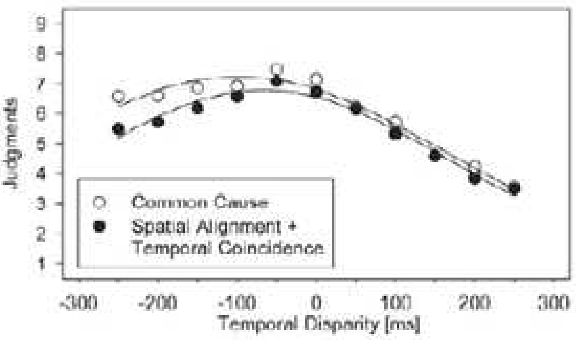


*Figure 2-2: Figure borrowed from Lewald & Guski (2003), showing that the temporal profile of inferring a common cause is slightly larger than the temporal profile of perceived synchrony for simple beeps and flashes.*

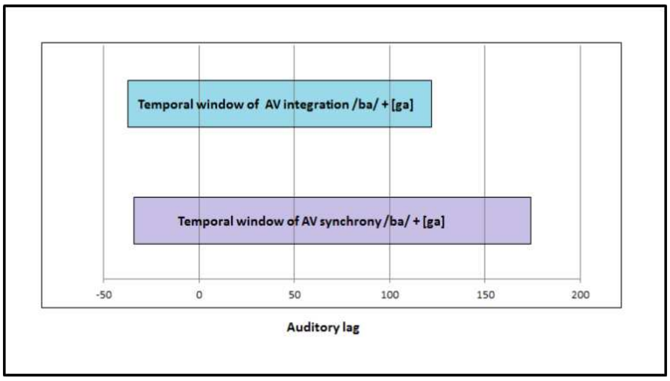


*Figure 2-3: Temporal windows of AV integration (blue) and perceived synchrony (purple), for incongruent combinations of the AV speech stimuli visual [ga] and auditory /ba / from van Wassenhove et al. (2007). The window of synchrony is larger than that of AV integration.*

perceived to be synchronous. Lewald and Guski's finding that participants attributed a common cause to stimuli they perceived as asynchronous are not entirely what would be predicted under the unity assumption. It is however possible that a change in response criteria across the two blocks may account for it. Furthermore, these temporal profiles were derived from grouped data, these measures were not directly compared against one another within participants and no correlational analysis was carried out. The asynchrony at which AV stimuli are most likely to be attributed to a common cause was also not obtained concurrently with the PSS, so whether information perceived as synchronous is at the same time also attributed to the same event is unclear.

Another account of AV integration which would argue for shared temporal mechanisms across implicit and explicit synchronisation is the automaticity account (illustrated in Figure 2.4 on the next page). According to this account, audiovisual integration is pre-attentive and automatic (Bertelson, Vroomen, de Gelder, & Driver, 2000) and results in no access to unisensory features once these are integrated into a multisensory percept, including features of the temporal relationship between the unisensory components. Under this premise, perceiving asynchrony is contingent upon *not* integrating AV information; once integrated, auditory and visual stimuli will be perceived as synchronous, even if they were physically asynchronous. Thus, according to the automaticity account, the temporal process underlying AV integration also determines the observer's explicit perception of AV relative timing, which would suggest that implicit and explicit AV synchronisation are performed by the same underlying mechanisms. This account would too predict a positive correlation between measures of ePSS and iPSS.

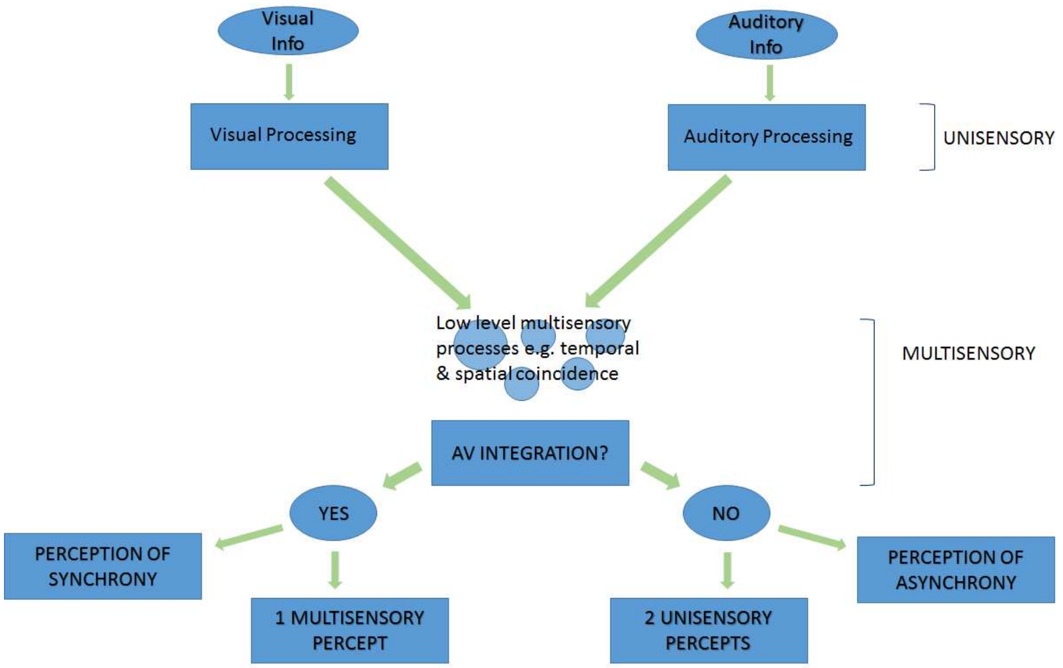


*Figure 2-4: Illustration of the automaticity account, whereby AV audiovisual integration is pre-attentive and automatic and results in no access to unisensory features once multisensory sensory information is integrated.*

### 2.1.2 THE CASE FOR MULTIPLE, DISTINCT SYNCHRONISATION MECHANISMS

Explicitly assessing the relative timing of visual and auditory stimuli is qualitatively different from making perceptual judgements required during tasks, such as the McGurk effect or the Stream-Bounce illusion described in Chapter 1, which measure audiovisual integration in the traditional sense. The different nature of the two tasks calls into question whether timing related measures derived from explicit timing tasks do indeed reflect the same mechanisms as those derived from implicit tasks and vice versa. After all, the perception of synchrony does not only occur for AV information which is readily integrated and incongruent auditory and visual stimuli which are unlikely to be integrated can still be perceived as synchronous (Vatakis, Ghazanfar, et al., 2008; Vatakis & Spence, 2008a). Moreover, AV stimuli which can be integrated such as typical McGurk AV syllables, are not always combined into an AV percept when perceived to be synchronous (Martin et al., 2012), thus AV integration and AV synchrony perception are not entirely dependent on one another.

Perceiving AV synchrony without necessarily integrating AV information is probably not uncommon, as temporal coincidence is not the only factor that plays a role in AV integration; for example attention (Alsius, Navarra, Campbell, & Soto-Faraco, 2005), semantic congruency (Vatakis & Spence, 2007) as well as prior knowledge (Petrini et al., 2009) are also contributing factors. What is perhaps more interesting is that auditory and visual information can sometimes be integrated when the visual and auditory streams are perceived to have occurred at different times (Soto-Faraco & Alsius, 2009; Soto-Faraco & Alsius, 2007). Using a dual-task paradigm, Soto-Faraco and colleagues measured perceived AV synchrony together with AV integration as a function of AV asynchrony within the same trials. The authors showed for the first time that on average, individuals can still integrate auditory and visual information presented at AV asynchronies which they can reliably detect as asynchronous at the same time. Over two experiments, they measured explicit synchrony perception using TOJs (Soto-Faraco & Alsius, 2007) and SJs (Soto-Faraco & Alsius, 2009), each concurrently with audiovisual integration using the McGurk illusion. On every trial participants made a phoneme identification judgement as well as a temporal judgement. For TOJs, their results showed (See Figure 2.5) that for SOAs between -160 and -400, illusory responses averaged at 42% whilst the auditory and visual components were being judged sound-first or sound-second on average 90% of the time. Similar results were found in the experiment which used SJs to measure perception of synchrony

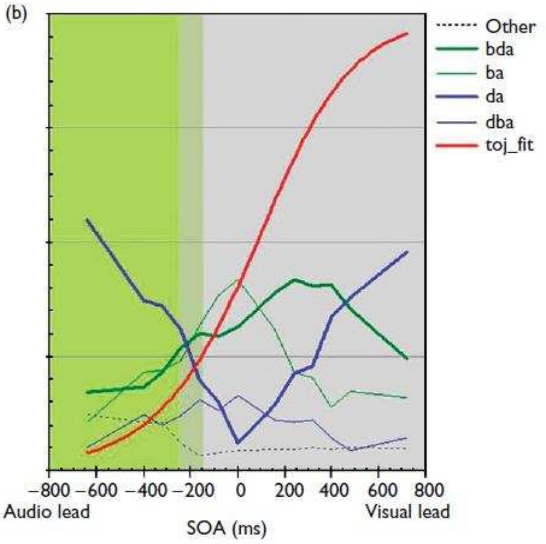


*Figure 2-5: Figure borrowed from Soto-Faraco & Alsius (2007), depicting temporal order judgement data superimposed onto Phoneme identification of McGurk AV stimuli. Illusory responses are observed at asynchronies which are reliably judged as either auditory leading or auditory lagging.*

(see Figure 2-6). These findings are indicative of a conflict between implicit and explicit synchrony percepts, which is contrary to what automaticity and unity accounts would predict. These results indicate that AV integration mechanisms seem to have judged AV stimuli as a common event, as the auditory and visual components were integrated to give rise to the McGurk illusion. At the same time, explicit synchronisation mechanisms judged the same AV stimuli as two distinct, asynchronous events. This is evidenced by the ‘asynchronous’ responses given by participants and their ability to discriminate AV temporal order reliably. In other words, participants experienced a concurrent “*dual perception*” (Salvador Soto-Faraco & Alsius, 2007, p348) regarding the timing of the stimuli pairs, which differed depending on which task they were performing. That is, participants perceived AV stimuli as both asynchronous and as a unified percept. This effect was however rather small, obtained by averaging data across participants, and could also reflect different response criteria across the two tasks.

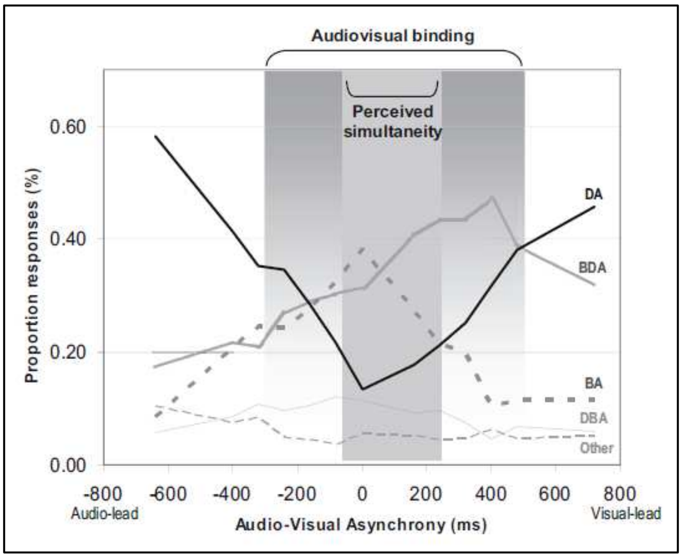


*Figure 2-6: Figure borrowed from Soto-Faraco & Alsius (2009), depicting synchrony judgement data superimposed onto Phoneme identification of McGurk AV stimuli. Illusory responses are observed at asynchronies which are reliably judged as asynchronous.*

A more extreme example of disunity of timing estimates across different audiovisual processes is the case of PH (Freeman et al., 2013), an otherwise normal individual who, following lesions in pons and basal ganglia, began to experience voices leading lips when watching people speaking. PH was tested on a temporal order dual-task paradigm, borrowed from Soto Faraco and colleagues (2007). As shown in **Error! Reference source not found.** on the next page, PH needed that lip-movements preceded the auditory stream by approximately 200ms in order to perceive the streams as synchronous. In contrast, in order to show maximal AV integration of auditory and visual stimuli, he needed lip-movements to *lag* the voice

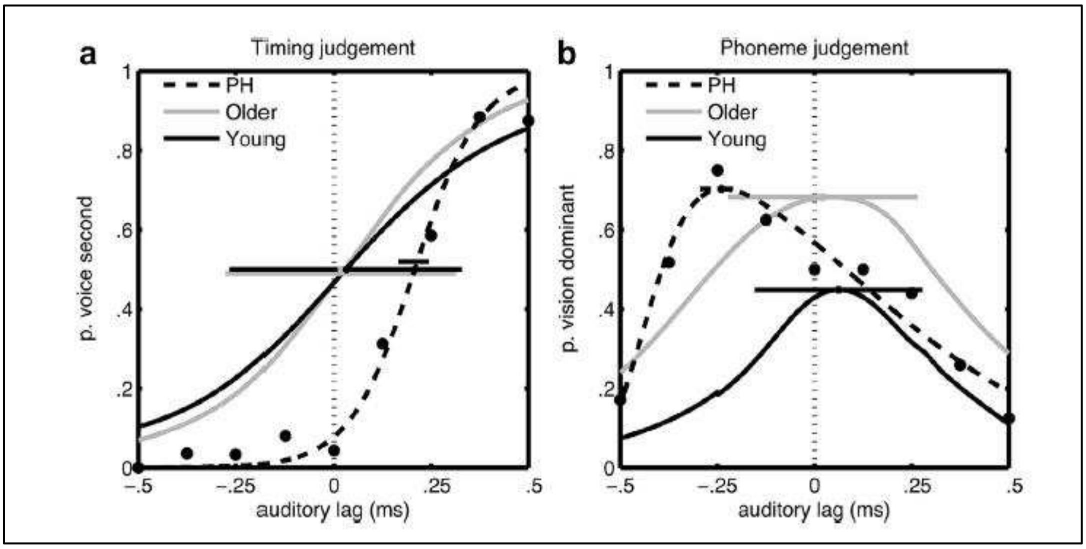


*Figure 2-7: Psychometric data for PH (black data points and broken line for psychometric fit), and controls (young: black continuous and older: grey). a) TOJ. b) Phoneme discrimination task, from Freeman et al. (2013).*

by the same amount of time (Freeman et al., 2013). The pons and basal ganglia have been reported to have pathways projecting to the auditory cortex (Halverson & Freeman, 2010; Kolomiets et al., 2001), so the location of PH's lesions suggest a slowing of propagation of auditory information, which explains the need to present auditory information before the onset of the visual stream for optimal AV integration. His lesion does not however account for his opposite explicit experience of voices leading lip-movements. Case studies have associated basal ganglia lesions with impairments in temporal sensitivity or duration perception (Grondin, 2010) so his lesion is also appropriate for disruption of time processing in general (Ivry & Spencer, 2004). However, with JND measures comparable to those of controls, PH seemed able to dissociate asynchronous from synchronous stimuli rather well around his PSS, and did not exhibit problems with duration perception (Freeman et al., 2013).

The case of PH motivated the current study, as his apparent fragmentation of AV timing estimates across implicit and explicit AV synchronisation is difficult to explain using existing accounts of AV synchronisation and integration. Amongst these are the notion that the brain strives to achieve unity as well as accounts of temporal ventriloquism (Aschersleben & Bertelson, 2003) or recalibration of temporal codes (Fujisaki et al., 2004; Hanson et al., 2008; Spence & Squire, 2003). If the brain achieved unity across different multisensory processes via these putative mechanisms, a consistent auditory delay in one mechanism would result in AV timing estimates of unaffected mechanisms being attracted towards this lag. That is to say, under such accounts, PH should exhibit a subjective auditory lag (or led) across *all* AV processes. Furthermore, most AV recalibration accounts posit that estimates of subjective timing are shifted towards constant auditory or visual lags to which the observer is exposed to, in order to account for factors such as observer-stimulus distance which lead to variability in the relative arrival time of AV information at the senses. According to such accounts, after some time, PH's experience of an auditory lag should go unnoticed as a result of cumulative adaptation towards it. Instead, PH's estimates of subjective AV synchrony derived from the temporal order and integration tasks are very different and symmetrically positioned on opposite sides of veridical synchrony, making him on average correct about the relative timing between auditory and visual stimuli across the two processes, but incorrect within each one.

Although there is some evidence to suggest that implicit and explicit synchronisation may not be reliant upon the same underlying mechanisms (Soto-Faraco & Alsius, 2009; Soto-Faraco & Alsius, 2007), this evidence is based on group averages and the effects are relatively small. It is commonly accepted that PSS measures are subject to individual differences (Boenke et al., 2009; Fujisaki et al., 2004b; van Eijk et al., 2010). If AV integration and explicit temporal judgements are indeed reliant on distinct temporal mechanisms (a possibility illustrated in Figure 2.8 on the next page), examining how individual variation in the iPSS relates to differences in the ePSS is more likely to reveal subtle differences between them, which may be concealed by group averages.

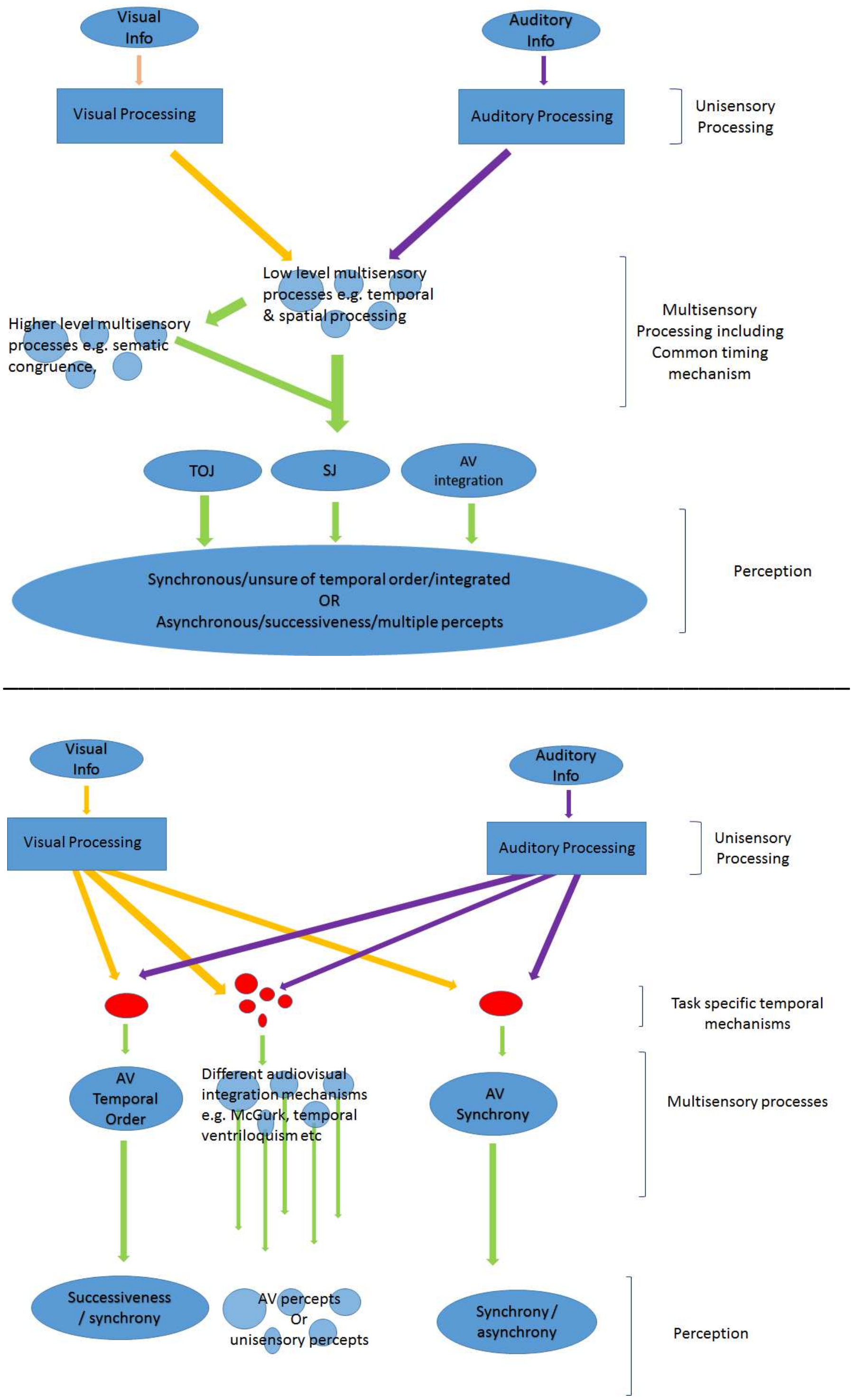


*Figure 2-8: Alternative underlying mechanisms of audiovisual timing. Top: potential underlying cognitive process contributing to SJs, AV integration and TOJs and resulting percepts if the tasks were supported by a common timing mechanism. Bottom: underlying processes contributing to SJs, TOJs and AV integration if these were served by multiple, task specific timing mechanisms.*

Alternatively, examining correlations between the iPSS and ePSS might reveal that the two correlate positively and that differences observed on average are likely to be the result of response bias or changes in decision criteria across different judgements. This would suggest that AV information does need to be perceived as synchronous to be integrated or vice versa. To date, no other study has attempted to examine the relationship between implicit and explicit synchrony estimates of subjective synchrony within participants using correlational analysis, which is a suitable approach for revealing whether measures of implicit and explicit subjective AV synchrony are statistically independent or dependent.

The following experiments address the question of whether subjective synchrony and audiovisual integration are likely to be supported by common mechanisms by examining the relationship between the asynchrony at which audiovisual integration is maximal, measured using the McGurk effect and the Stream bounce illusion, and PSS measures derived using TOJ and SJ. Subjective synchrony and audiovisual integration were concurrently measured as a function of AV asynchrony using a dual-task paradigm borrowed from Soto-Faraco and Alsius (2007, 2009). If unity is typically achieved across mechanisms supporting subjective synchrony and AV integration (a scenario seen on the top part of Figure 2.8, on the previous page), then sound and vision should be optimally integrated when they are perceived to be synchronous most often. In other words, under the assumptions of unity and automaticity we expected the point of subjective synchrony (PSS) to be positively correlated with the AV asynchrony optimal for the McGurk effect. On the other hand, if these measures do not correlate positively, it might be an indication that unity is not achieved across different audiovisual mechanisms.

# 2.2 EXPERIMENT 2.1: MCGURK

## 2.2.1 METHODS

### 2.2.1.1 SUBJECTS

Twenty-seven neurologically healthy young subjects (18-28 years, mean 22) took part in the experiment. Data from four other participants were excluded, due to poor performance, resulting in implausible estimates of subjective timing >300ms asynchrony, outside the typical range for multisensory integration (Munhall et al., 1996; van Wassenhove et al., 2007; Vatakis, Ghazanfar, et al., 2008; Vatakis & Spence, 2007) and indicative of poor quality data and unreliable function fits.

### 2.2.1.2 STIMULI AND APPARATUS

Laboratory apparatus comprised an Apple Mac Mini, with Labtech speakers positioned either side of a 17" Sony HMD-A420 CRT display, viewed in darkness from 70cm. Video mode was 1200 x 800 with a 85Hz refresh rate. Subjects responded using the cursor keys on a standard keyboard.

McGurk stimuli were based on Soto-Faraco & Alsius (2007), which were kindly provided by the authors. Auditory /ba/ and /da/ phonemes (with white noise at 15% of maximum amplitude) were combined with visual lip-movements for [ba] and [ga]. The two incongruent pairings for eliciting the McGurk effect were /ba/+[ga]=’da’ and /da/+[ba]=’ba’ or ‘bda’. The other two ‘congruent’ pairings /ba/+[ba] and /da/+[da] tend to be heard correctly. Background was set to the average RGB value across all pixels and frames.

### 2.2.1.3 DESIGN

The experiments employed a repeated measures factorial design. Audiovisual asynchrony was manipulated so that the soundtrack was presented at each of 9 auditory lags relative to the visual sequence including synchronous (0ms) presentation, within a range of ±500ms. A second independent variable was the congruency of lip-movements with voice (see Stimuli section above). There were two possible lip-voice combinations for each congruent/incongruent pairing. Only incongruous conditions were used for assessing McGurk interference. Two sets of dependent measures were obtained from two responses elicited after each trial, for temporal judgments and phoneme identity. In one condition participants performed phoneme judgements concurrently with TOJs in another they performed phoneme judgements with SJs.

### 2.2.1.4 PROCEDURE

A dual-task paradigm (Soto-Faraco & Alsius, 2007) (see Figure 2-9 on the next page) was used to obtain the measures. The experiment was carried out in a dark laboratory. Before each condition, participants were familiarized with the tasks, and given a practice block to complete. A fixation display was presented at the beginning of each trial. Participants were required to press the space bar when ready to begin each trial. After a randomly selected delay of 1000±500ms, an audiovisual clip was displayed for 2800ms. On each trial, audiovisual asynchrony and stimulus pairing were selected pseudo-randomly. In each condition, each of nine possible asynchronies was presented for each of the stimulus pairing and repeated a minimum of 8 times, giving a minimum of 288 trials. Following movie offset, two successive forced-choice questions were displayed on the screen. In the TOJ condition, participants were first required to judge the temporal order of the stimuli, by stating whether the voice onset preceded or followed the lip-movement. After their response, participants were then asked to discriminate the spoken phoneme by reporting the whether they heard “ba” or “da” (a third option for ‘other’, used on only 0.3% ± 0.3%SEM of trials, was not included in further analysis). Subjects were encouraged to choose the option that sounded the closest to what they heard. In the SJ condition, participants first performed a synchrony judgement on the stimuli by stating whether the voice and lip-movements were presented synchronously or asynchronously, after which they performed the phoneme identification described above. The order of SJ and TOJ blocks were counterbalanced across participants.

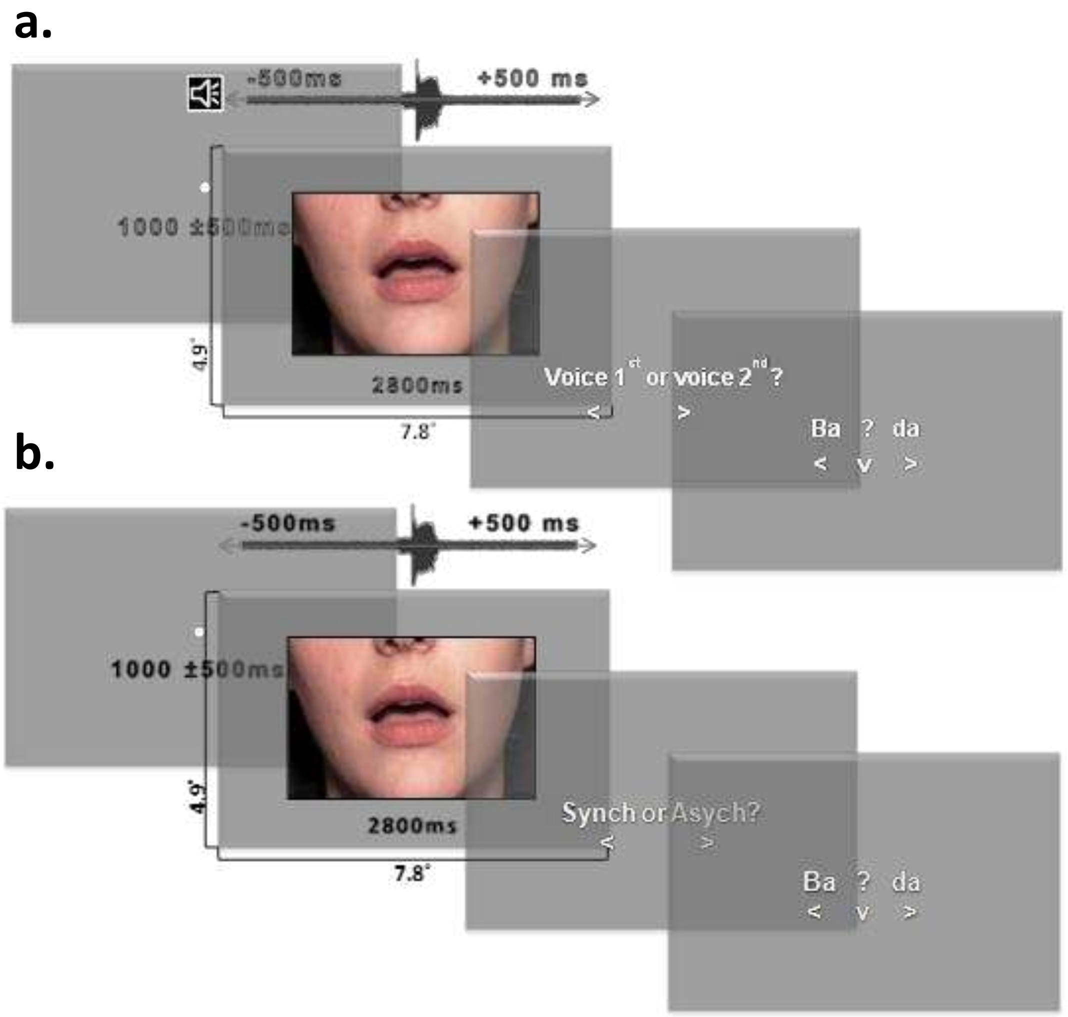


*Figure 2-9: Trial sequence and stimuli for McGurk* ***a.*** *TOJ dual-task and* ***b.*** *SJ dual task.*

### 2.2.1.5 ANALYSIS

#### 2.2.1.5.1 Temporal order judgements (TOJ)

For TOJ, the proportion of ‘voice second’ responses (where the auditory onset was judged to lag the visual onset) was plotted as a psychometric function of actual auditory lag time in milliseconds (note that negative lag denotes an auditory lead). The proportion of ‘sound second’ values was typically below 50% for negative auditory lags (i.e., sound leads vision), and above 50% for positive auditory lags. A logistic function was then fitted to the psychometric data, using a maximum-likelihood algorithm provided by the PSIGNIFIT toolbox for Matlab (Wichmann and Hill, 2001). The critical auditory lag corresponding to the participant's PSS was then read off from the fitted function. This is the point at which the participant is at chance (50%) deciding whether the sound came first or second relative to the visual onset. The same software was used to find the slope of the function and to derive 95% confidence intervals for both PSS and slope estimates, via a bootstrapping procedure. Finally, additional auditory lag required for the participant to switch from responding at chance to responding 'voice second' 75% of the time was estimated. The resulting value quantifies the lag that can produce a Just Noticeable Difference (JND) between subjectively synchronous and asynchronous stimuli.

The above procedure was carried out for each of the four audiovisual conditions (2 congruent and 2 incongruent) as well as for the average proportions across all the conditions (see Figure 2.10 below). In order to obtain the most representative PSS and JND measures, the 5 resulting measures of PSS and of JND were averaged to obtain a final measure of each parameter.

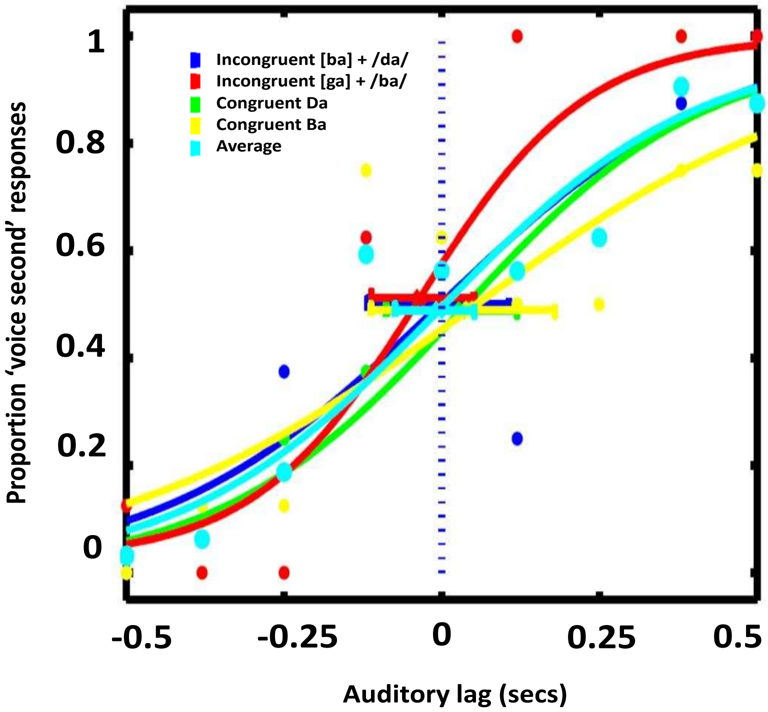


*Figure 2-10: Raw TOJ data (proportion of 'voice second' responses) from the 4 audiovisual conditions as well as their average (dots) plotted as a function of audiovisual lag with psychometric functions fitted to data.*

### 2.2.1.5.2 Synchrony Judgement (SJ)

For the SJ task, the proportion of 'synchronous' responses was plotted as a function of asynchrony. This function was then fitted with a symmetrical Gaussian function, using the fminsearchbnd function in Matlab. This implemented an iterative algorithm for finding the parameters of the function that minimised the residuals given the raw data. The PSS was then read off from the fitted psychometric function as the auditory lag which corresponded with the peak of the Gaussian. The standard deviation of the function provided an estimate of participants' temporal precision (SD) in synchrony judgements.

The above procedure was carried out for each of the four audiovisual conditions (2 congruent and 2 incongruent) as well as for the average proportions across all the conditions (see Figure 2.11 below)). In order to obtain the most representative PSS and temporal precision measures, the 5 resulting measures of PSS and of SD were averaged to obtain a final measure of each parameter.

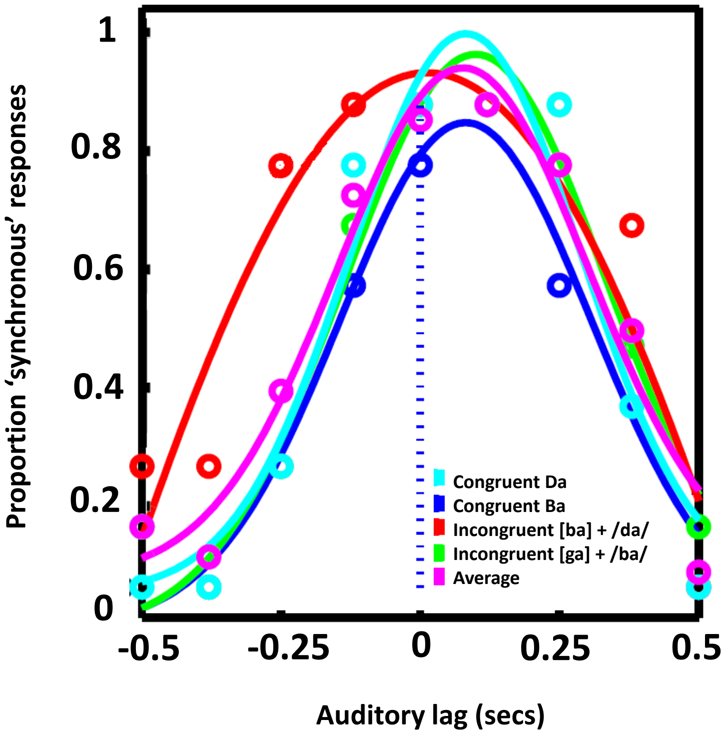


*Figure 2-11: Raw SJ data (proportion of 'synchronous' responses) from the 4 audiovisual conditions as well as their average (circles) plotted as a function of audiovisual lag with psychometric functions fitted to data.*

### 2.2.1.5.3 Phoneme identification (McGurk)

For the phoneme discrimination task, the proportion of trials in which the reported phoneme was consistent with the lip-movements was obtained, averaged across incongruous conditions only. For example, a ‘ba’ response to /da/ + [ba] and a ‘da’ response to /ba/ + [ga] were scored as ‘consistent’. This was plotted as a psychometric function of auditory lag. The data from each of the two incongruent conditions, plus their average, were fit using an asymmetric double sigmoid function (ADS, following van Wassenhove et al., 2007), which results in a bell-shaped curve with adjustable height, width and asymmetry, using the following equation:

$$y = \frac{1}{2}\left[tanh\left(\frac{x - c_1}{w_1}\right) - tanh\left(\frac{x - c_2}{w_2}\right)\right]$$

With constraints $w_1>0$ and $w_2>0$

The optimal auditory lag for maximum McGurk interference (implicit PSS; iPSS) was read off at the peak of each of these interpolated functions and averaged, with 95% confidence intervals derived from fits of 1000 bootstrapped samples.

The above procedure was carried out for each of the two incongruent audiovisual conditions as well as for the average proportions across them (see Figure 2.12 on the next page). In order to obtain the most representative measures, the 3 resulting measures of iPSS were averaged to obtain a final measure.

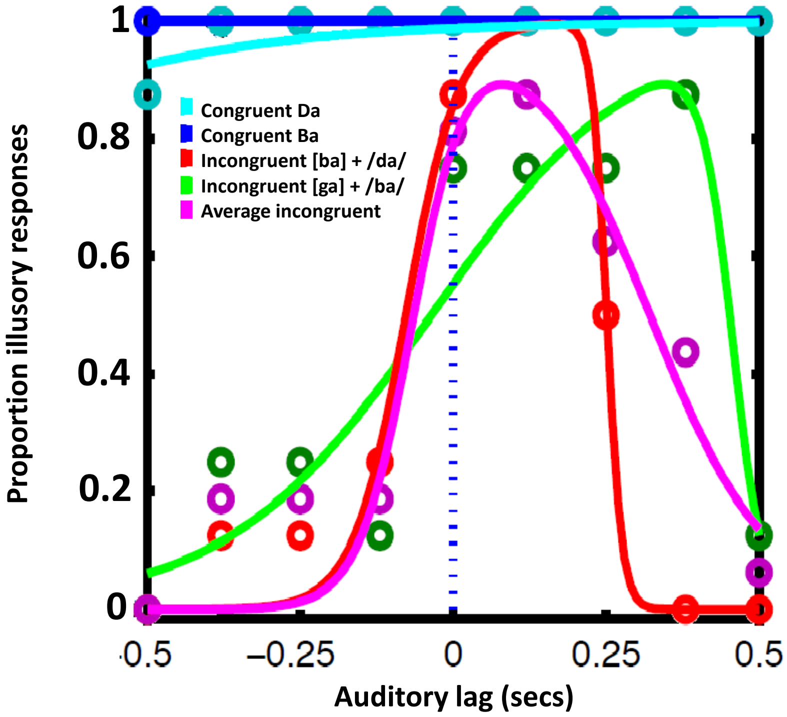


*Figure 2-12: Raw McGurk proportion (proportions of visually driven responses) plotted as a function of auditory lag, with ADS functions fitted to the data. Only the parameters for the incongruent conditions and from their average were used to calculate the final measures.*

## 2.2.2 Results

### 2.2.2.1 TOJ ePSS and McG iPSS

Both TOJ ePSS data were normally distributed (statistics in table 2-1, below), but the McGurk iPSS data were not. A non-parametric correlation was carried out. A significant negative correlation (illustrated in Figure 2.13 below) was found between measures of ePSS derived from TOJs and measures of iPSS derived from the McGurk effect [$r(27)= -.519$, $p=.006$].

| ***Measure*** | *Shapiro Wilk statistics* |
|---|---|
| ***TOJ ePSS*** | $W(27)=.96$, $p=.381$ |
| ***McGurk iPSS*** | $W(27)=.91$, $p=.027$ |

*Table 2-1: Shapiro Wilk statistics for normality of distribution of TOJ ePSS and McG iPSS*

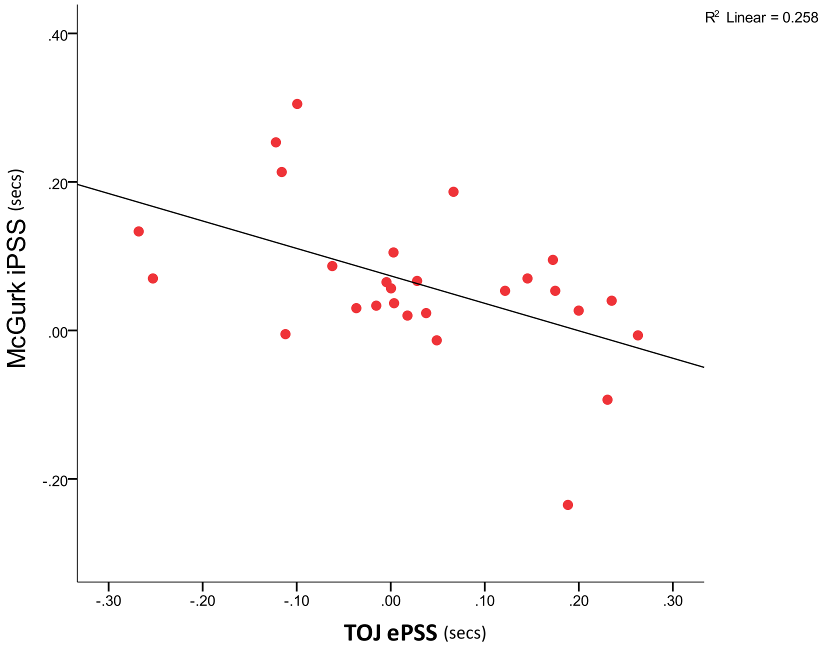


*Figure 2-13: Scatter plot of significant negative correlation between McG iPSS and TOJ ePSS*

### 2.2.2.2 SJ ePSS and McGurk iPSS

One data point was excluded from the analysis due to a flat phoneme identification function. The SJ PSS data and McGurk iPSS data were normally distributed (see table 2-2 below for statistics). A parametric correlation was run (illustrated in Figure 2.14 below). There was no significant correlation between measures of McGurk iPSS and the ePSS derived from SJs [*r*(26)=.215*, p*=.303].

| ***Measure*** | *Shapiro Wilk statistics* |
|---|---|
| ***SJ ePSS*** | *W*(26)=.96, *p*=.326 |
| ***McGurk iPSS*** | *W*(26)=.95, *p*=.242 |

*Table 2-2: Shapiro Wilk statistics for normality of distribution of SJ ePSS and McG iPSS*

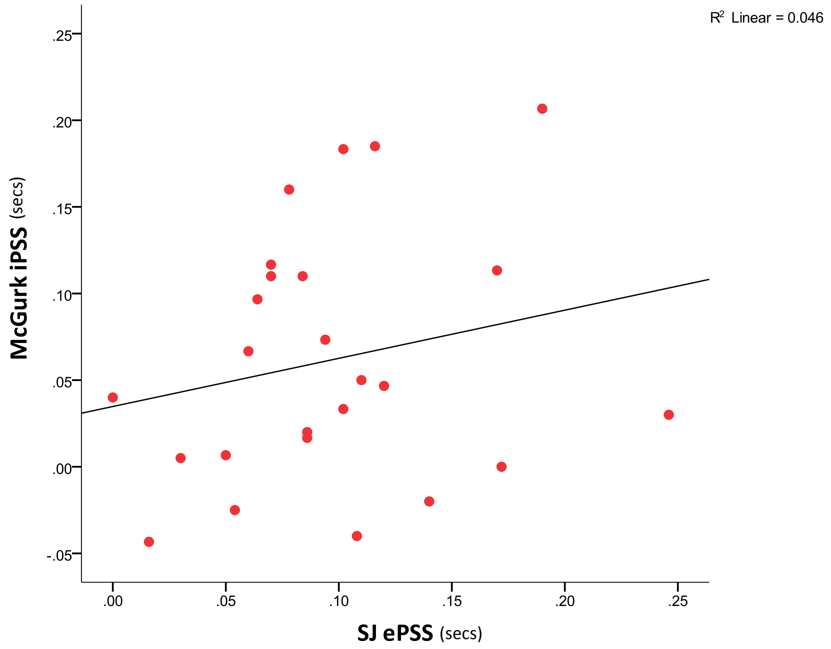


*Figure 2-14: Scatter plot of non-significant negative correlation between McG iPSS and SJ ePSS*

### 2.2.2.3 SJ EPSS AND TOJ EPSS

A further analysis was carried out to examine the correlation between TOJ ePSS and SJ ePSS, illustrated in Figure 2.15, below. There was no significant correlation between these measures [*r*(27) =.274, *p*=.167].

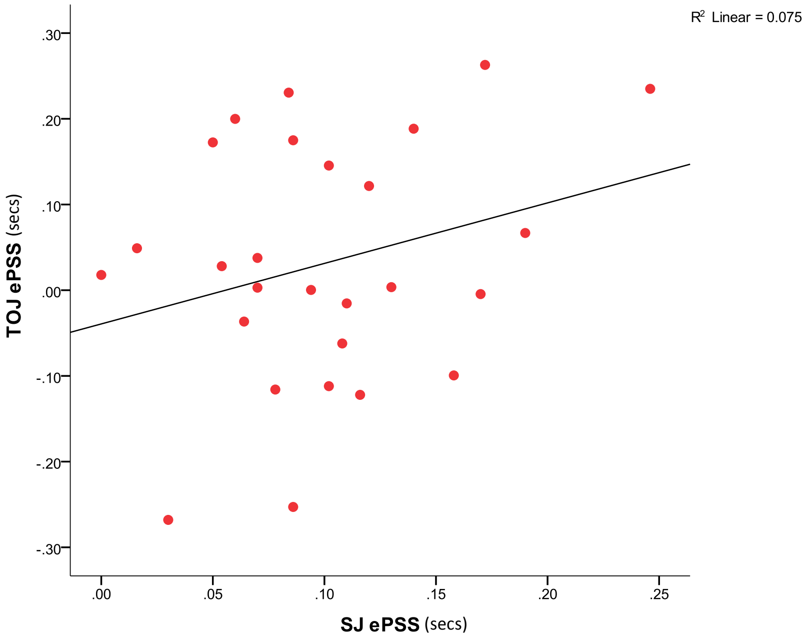


*Figure 2-15: Scatter plot of (non-significant) correlation between TOJ ePSS and SJ ePSS*

# 2.3 EXPERIMENT 2.2: STREAM BOUNCE

A good way to check the generalizability of the results obtained in Experiment 2.1 is to replicate the dual-task paradigm using non-speech stimuli. The Stream-Bounce illusion (Sekuler et al., 1997) is suitable for this, for the following reasons. Like the McGurk illusion used in Experiment 2.1, it allows for the concurrent measurement of ePSS and iPSS. The visual component of the Stream-Bounce illusion is dynamic, and its duration is similar to that of the visual speech stimuli. The auditory and visual components of the Stream-Bounce display can be presented at the same SOAs as those of the McGurk stimulus. Furthermore, the direction of influence in the Stream-Bounce illusion is from sound to vision, whereas in the McGurk effect it is from vision to sound, making the replication a good test of generality. To this end, the dual-task paradigm was replicated using the Stream-Bounce illusion in Experiment 2.2.

## 2.3.1 METHODS

### 2.3.1.1 SUBJECTS

Twenty eight new participants (2 males) aged 18-24 took part in the experiment. The data from two others were excluded as they did not report perceiving the disks bouncing on any of the trials. All participants were naïve to the specific aims of this study. Participants received course credits amounting to the length of time they spent completing the tasks. Procedures were approved by the local Psychology ethics committee.

### 2.3.1.2 STIMULI AND APPARATUS

The same laboratory apparatus was used in this experiment as in Experiment 2.1. Visual stimuli were two yellow circular disks at maximum contrast on a black background. Each moved from positions left and right above fixation, via the central fixation point where they would meet, and then continue moving to opposite positions below fixation (see Figure 2.16 for dimensions in the ‘Procedure’ section).

Animations were accompanied by a 400Hz tone of 100ms duration. Movies were followed by 9pt white text prompting responses, displayed centrally.

#### 2.3.1.3 DESIGN

The experiments employed a repeated measures factorial design. The audiovisual asynchrony was manipulated, so that the soundtrack could be shifted forwards or backwards in time relative to the collision point of the two disks, over a range of ±500ms through nine equal steps of 125ms including zero (sound synchronous with the collision point). Each asynchrony condition was presented 20 times resulting in 180 trials. In the TOJ condition, participants made a TOJ followed by a Stream-Bounce judgement and in the SJ condition they made a SJ judgement, also followed by a Stream-Bounce judgement (see Figure 2-16 on the next page).

#### 2.3.1.4 PROCEDURE

A dual-task paradigm (Soto-Faraco & Alsius, 2007) (see on next page) to obtain two concurrent measures of the audiovisual asynchrony that is (1) perceived as synchronous, and (2) optimal for audiovisual integration. Experiments were carried out in a dark laboratory. Participants were familiarised with the task and given a practice block of 30 trials before starting each condition. On each trial, participants pressed the space bar on a standard keyboard when they were ready to view the stimuli. After the AV display, in the TOJ condition participants made a temporal order judgement by indicating whether the beep occurred before or after the disks reached the middle point of the screen, followed by a judgement on whether the disks appeared to stream through, or bounce off each other. In the SJ condition, on each trial participants made a synchrony judgement by indicating whether the beep occurred at the same time or at a different time to the point at which the disks reached the middle of the display, followed by the stream/bounce judgement. The order in which the temporal judgement conditions were carried out was counterbalanced across participants.

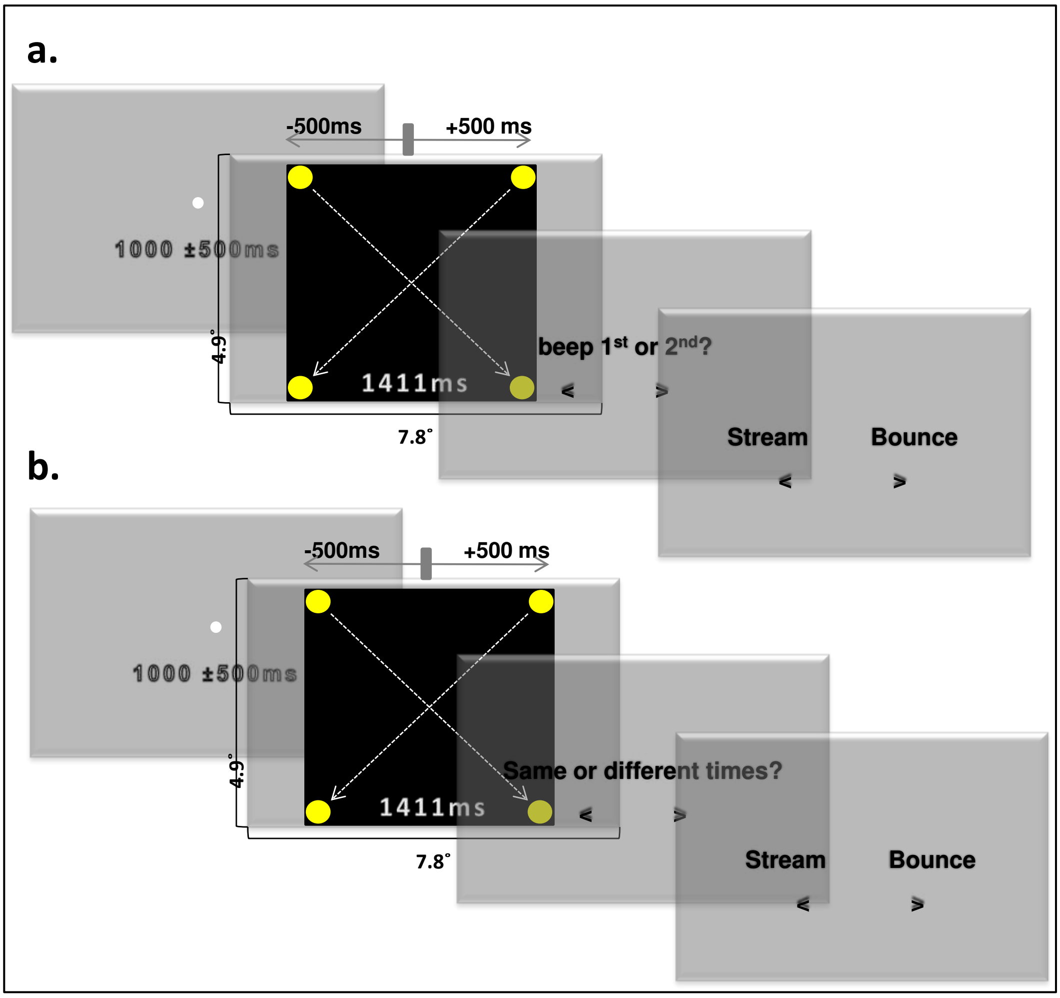


*Figure 2-16: Trial sequence and stimuli for the Stream Bounce* ***a. TOJ*** *and* ***b. SJ*** *dual tasks*

### 2.3.1.5 ANALYSIS

Data were split into two, and fitted using the same procedures used in the previous experiment (see Section 2.1). For the stream-Bounce illusion, 'bounce' responses were plotted as a function of Auditory lag and handled in the same way as McGurk responses were plotted in the previous section. Examples of the fitting procedures can be seen in Figures 2-17, 2-18 and 2-19 on the next two pages.

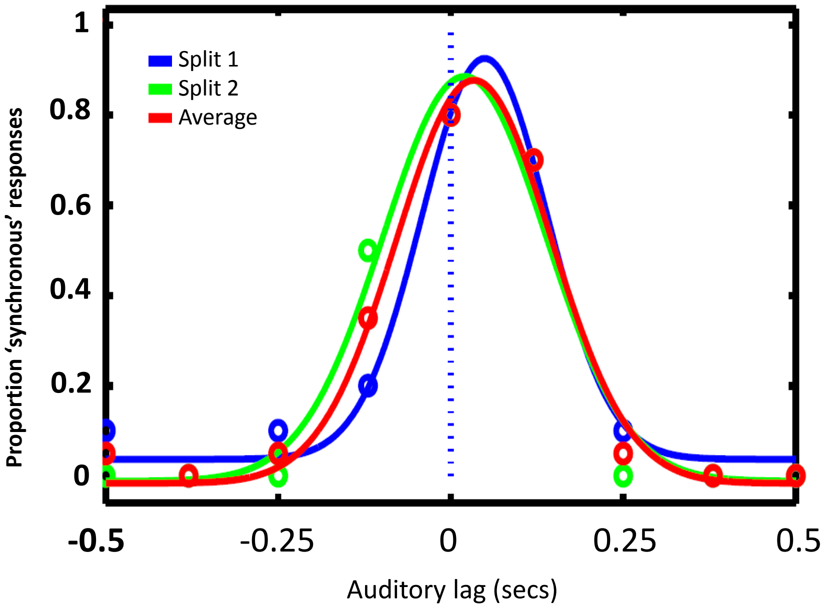


*Figure 2-17: Example of fitting procedure for synchrony judgement data. Raw data split into two (proportion of 'synchronous' responses – circles) as well as the average of the two halves plotted as a function of auditory lag, with symmetrical Gaussian curves fitted to the data.*

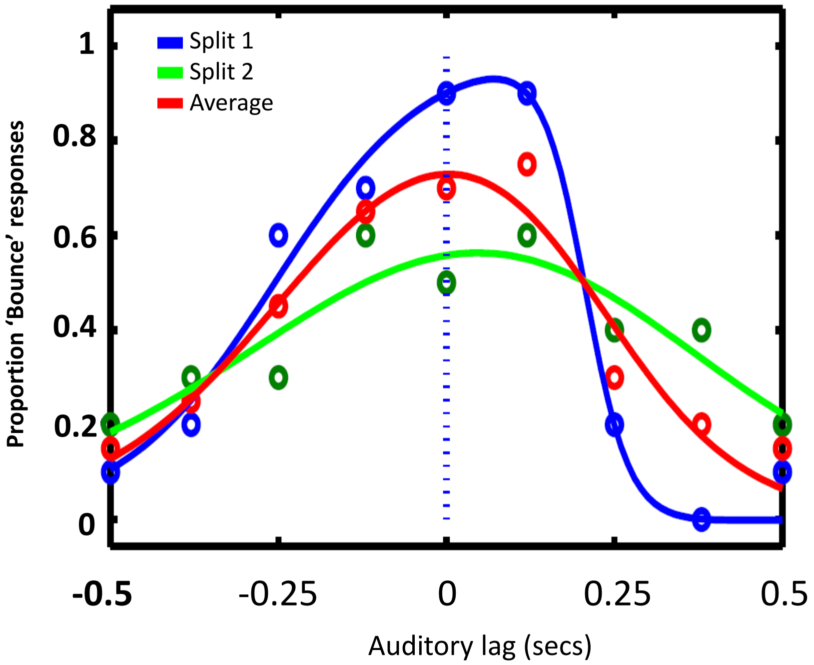


*Figure 2-18: Example of fitting procedure for Stream/Bounce data. Raw data was split into two and proportion of 'bounce' responses (circles) as well as the average of the two halves, were plotted as a function of auditory lag. ADS curves were fitted to the raw data, the peaks (iPSS) were estimated and averaged across the three.*

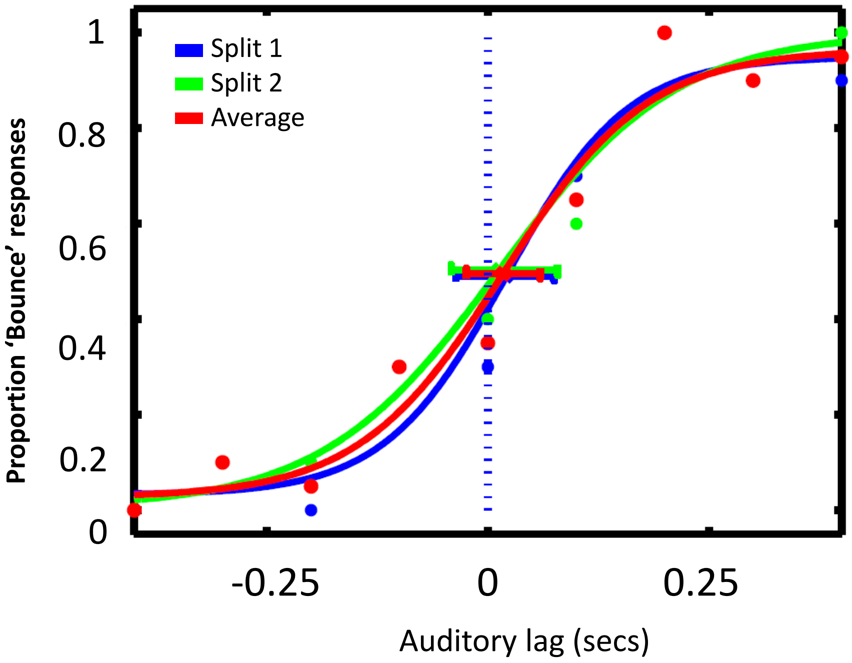


*Figure 2-19: Example of fitting procedure for TOJ data. Raw data was split into two and proportion of 'sound-second' responses (circles), as well as the average of the two halves, were plotted as a function of auditory lag. Logistic functions were fitted to the raw data, PSS measures were estimated and averaged across the three curves.*

## 2.3.2 RESULTS

### 2.3.2.1 TOJ ePSS AND STREAM-BOUNCE iPSS

The TOJ ePSS and Stream-Bounce iPSS data were normally distributed (Statistics in table 2-3, below). A parametric correlation was carried out. A significant, negative correlation (See Figure 2.20) was found between measures of the ePSS derived from TOJS and the iPSS derived from Stream-Bounce judgements [r(28)=-.425, *p*=.024].

| ***Measure*** | *Shapiro Wilk statistics* |
|---|---|
| ***TOJ ePSS*** | *W*(28)=.99, *p*=.941 |
| ***Stream Bounce iPSS*** | *W*(28)=.97, *p*=.669 |

*Table 2-3: Shapiro Wilk statistics for normality of distribution of TOJ ePSS and Stream-Bounce iPSS*

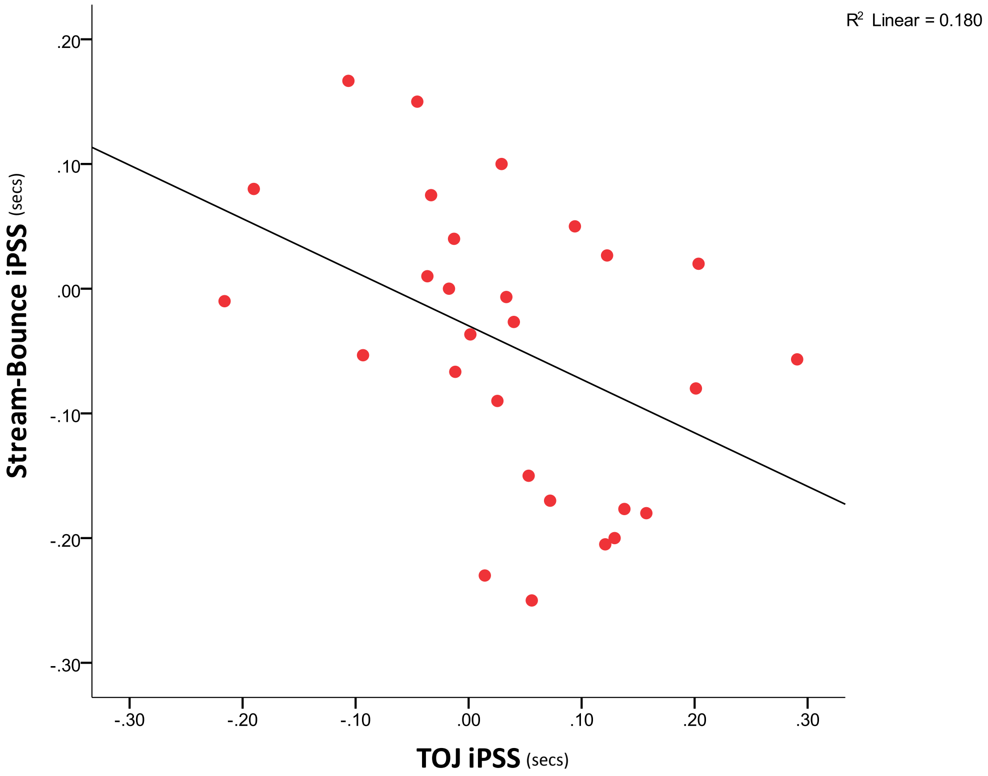


*Figure 2-20: Scatter plot of significant negative correlation between Stream-Bounce iPSS and TOJ ePSS*

### 2.3.2.2 SJ ePSS AND STREAM BOUNCE iPSS

Stream-Bounce iPSS measures were normally distributed, but SJ ePSS measures were not normally distributed (see table 2-4 below for statistics). A non-parametric correlation was carried out. Measures of Stream-Bounce iPSS correlated significantly and positively with measures of ePSS derived from SJs [*r*(28)=.547, *p*=.003] (see Figure 2.21). Closer inspection of the data however revealed that this significant correlation was driven by an outlier (>2SD from mean of ePSS and iPSS). Once this outlier was excluded from the analysis the correlation still had a positive trend, but was no longer significant at the 5% level of confidence [*r*(27)=.355, *p*=.069, two tailed] (See Figure 2.22, on the next page).

| ***Measure*** | *Shapiro Wilk statistics* |
|---|---|
| ***SJ ePSS*** | *W*(27)=.75, *p<.001* |
| ***Stream Bounce iPSS*** | *W*(27)=.97, *p*=.494 |

*Table 2-4: Shapiro Wilk statistics for normality of distribution of SJJ ePSS and Stream-Bounce iPSS*

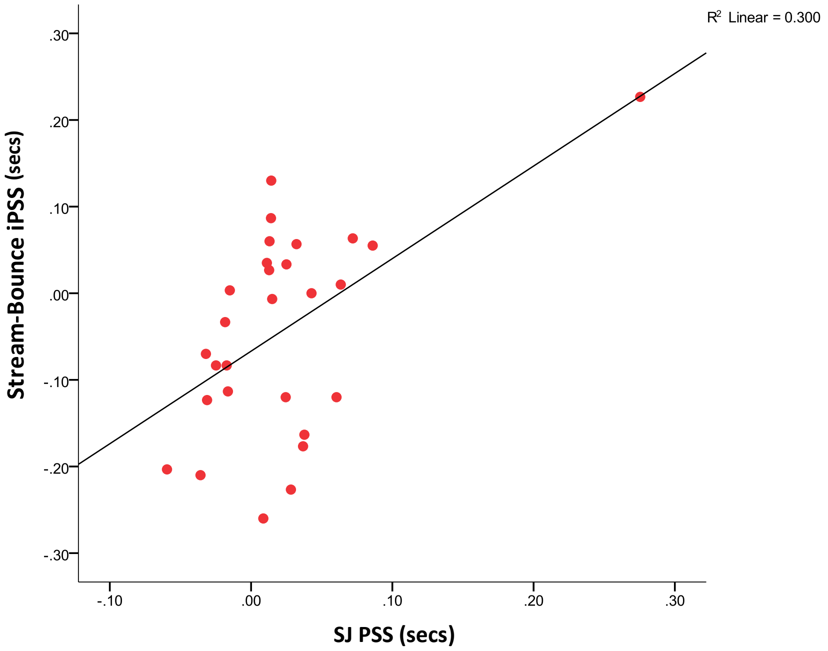


*Figure 2-21: Scatter plot of significant positive correlation between SJ ePSS and Stream-bounce iPSS, driven solely by one outlier*

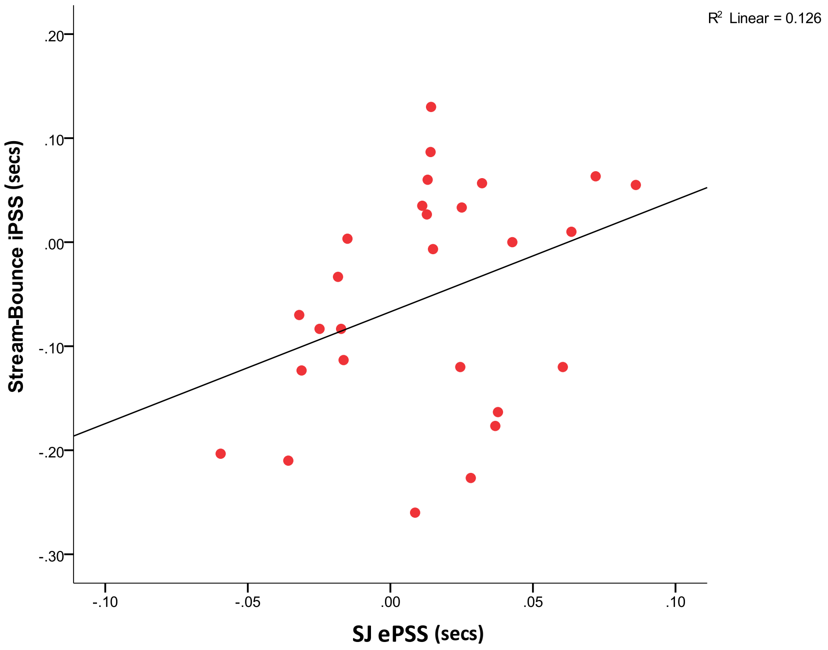


*Figure 2-22: Scatter plot of non-significant correlation between SJ ePSS and Stream-bounce iPSS, without the outlier*

In case this particular participant's data set had a similar effect on the correlation between TOJ ePSS and Stream-Bounce iPSS, an additional analysis was also carried out on this data, excluding this participant's data. The correlation remained negative and significant at the 5% level of confidence [*r*(27)=-.396, p=.042, two tailed].

### 2.3.2.3 TOJ EPSS AND SJ EPSS

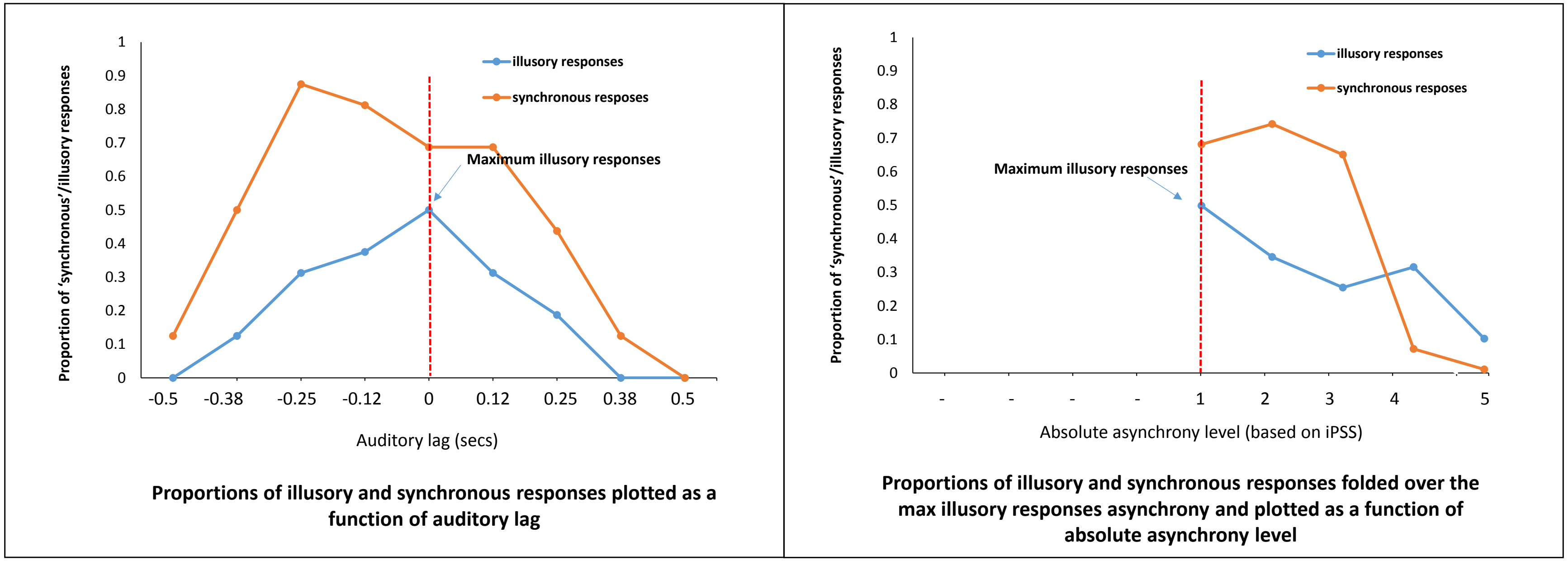
An additional analysis was run to examine whether estimates of TOJ ePSS and SJ ePSS were related to one another. There was no significant relationship between TOJ ePSS and SJ ePSS [*r*(27)=.282, *p*=.154].

## 2.4 SUBJECT-BY-SUBJECT ANALYSIS OF SJ DATA

As no significant correlation was found between iPSS and ePSS derived from SJs, the following section employs an alternative analysis to examine whether explicit synchrony responses can significantly predict the implicit (AV integration) responses on each trial, on a subject-by-subject basis, over and above AV asynchrony. If the explicit response does predict the implicit one over and above the asynchrony variable, it would provide evidence that the two responses were not made independently from one another, which would in turn suggest that the explicit response was influenced by the implicit one, or vice versa.

For each of the two SJ dual task data sets, logistic hierarchical regressions were carried out with the implicit response entered as the dependent variable. The absolute AV asynchrony was entered as the sole predictor in block 1 and the explicit response was added to the model in the second block. If variability in the implicit response is explained by the variability in the explicit response, over and above the variability explained by the change in absolute AV asynchrony, it would mean that explicit perception of synchrony might have an effect on whether AV information is integrated or not (or vice versa).

For McGurk data, only data from the incongruent conditions were used for the analysis. For Stream-Bounce data, all data were used. Raw data was plotted as a function of auditory lag (see Figure 2.23 on the next page for example of data processing from McGurk data) and the asynchrony variable was folded over the point at which AV integration occurred most frequently in order to obtain levels of absolute asynchrony centred around the optimal asynchrony for AV integration. AV asynchronies were then recoded 1 to ‘n’ (n being the number of resulting asynchrony levels) and entered as the first predictor in the regression. Responses were recoded as 1 for synchronous and illusory responses and 0 for asynchronous and veridical responses. The explicit response was then entered as the dependent variable and the implicit response as the additional predictor in block 2.



**Proportions of illusory and synchronous responses plotted as a function of auditory lag**

**Proportions of illusory and synchronous responses folded over the max illusory responses asynchrony and plotted as a function of absolute asynchrony level**

*Figure 2-23: Left: Proportion of illusory and synchronous responses plotted as a function of auditory lag (secs) in order to determine the asynchrony at which maximum illusory responses occur (red dotted line). Right: data folded over the point at which maximum illusory responses occur, and proportions of illusory and synchronous responses plotted as a function of absolute asynchrony (levels, each 0.12 secs apart). Once the optimal asynchrony was found, auditory lag was recoded 1 to 'n' (n being the number of resulting asynchrony levels. The data were then entered into a logistic regression with synchronous and illusory responses coded as 1 and asynchronous and veridical responses coded as 0.*

### 2.4.1 MCGURK

A total of 27 regressions were run. For 19 participants, the explicit response was non-significant and its addition did not make a significant improvement to the model. In other words, the implicit response could not be predicted from the explicit response over and above AV asynchrony. For a further 5 participants, the addition of the explicit response did make a significant improvement in the model and also rendered the asynchrony variable non-significant. For a further 2 subjects, the addition of the implicit response variable made a significant improvement in the model, but did not render the asynchrony variable non-significant. Both asynchrony and explicit response were significant predictors of the implicit response, each explaining unique variability in the implicit response variable. For one further participant, neither the asynchrony nor explicit response variables significantly predicted the implicit response.

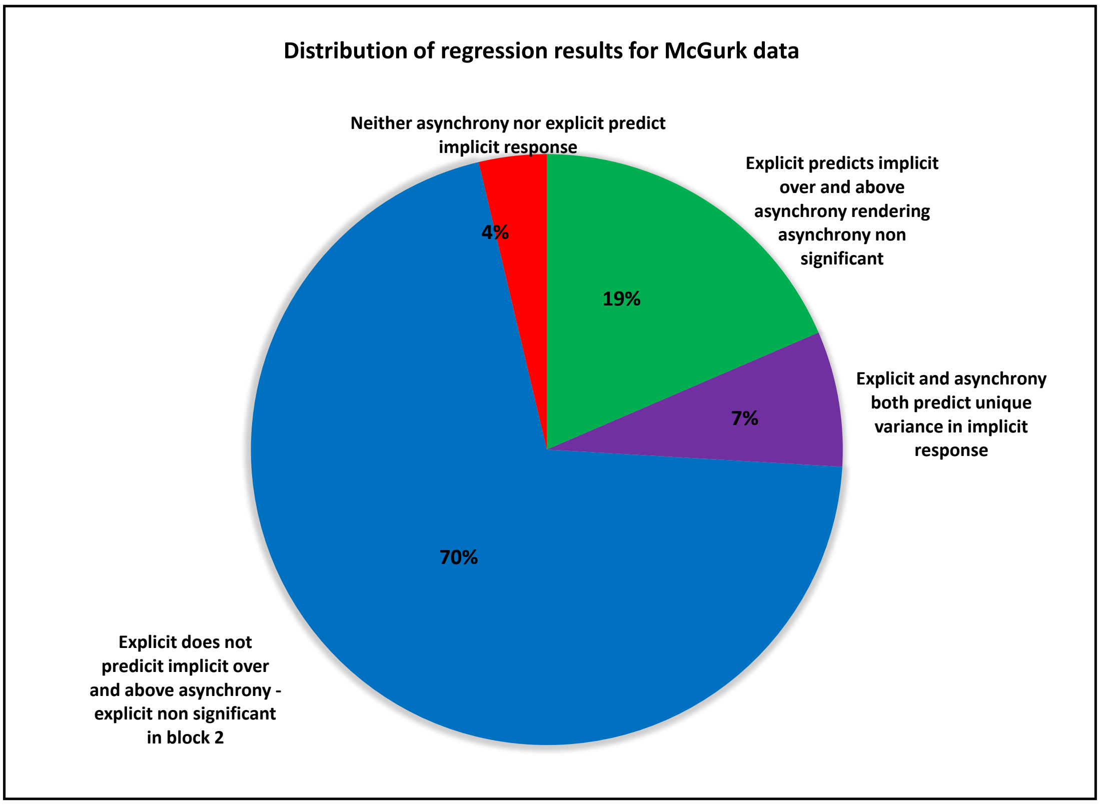


*Figure 2-24: Distribution of regression results for McGurk-SJ dual-task data.*

## 2.4.2 STREAM BOUNCE

Twenty-eight hierarchical regressions were run with the explicit response as the dependent variable, absolute asynchrony level as the predictor in block one and implicit response as the additional predictor in block two. The overall results are illustrated in a pie chart in Figure 2.25, below.

For 11 subjects, the explicit response significantly predicted the implicit response, over and above AV asynchrony. The addition of the implicit response resulted in the asynchrony variable being a non-significant predictor of the implicit response.

For a further participant, AV asynchrony did not predict the implicit response, but the explicit response did. For an additional participant, asynchrony did not significantly predict implicit response, and the addition of the explicit response led to a borderline significant improvement in the model (p=0.55) with the explicit response being a borderline significant (p=0.56) predictor of the implicit response.

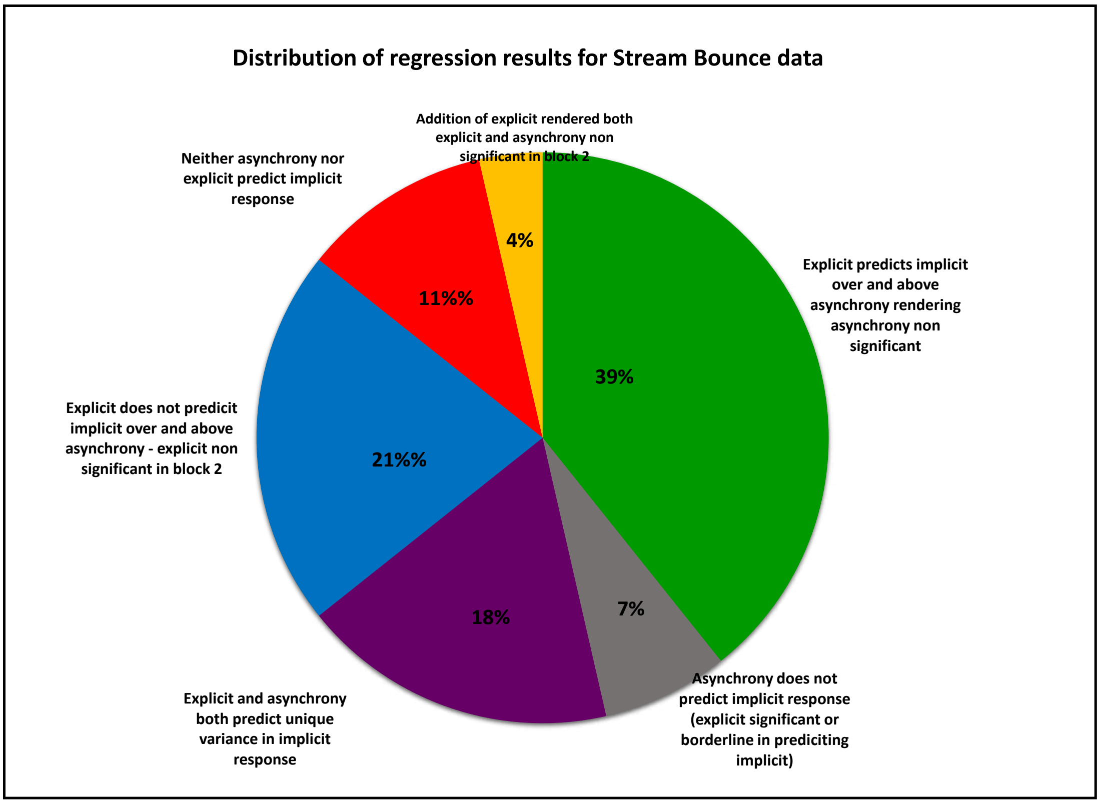


*Figure 2-25: Distribution of regression results for Stream-Bounce-SJ dual-task data.*

For 5 further participants, the addition of the implicit response variable made a significant improvement in the model, but did not render the asynchrony variable non-significant. Both asynchrony and explicit response were significant predictors of the implicit response, each explaining unique variability in the implicit response variable. For another 6 participants, the addition of the explicit response did not make a significant improvement in the model, leaving AV asynchrony as the only significant predictor of the implicit response. For an additional participant, the addition of the explicit response did make a significant improvement to the model, but the explicit response variable itself was non-significant within that new model. The asynchrony variable was still the only significant predictor in the model.

For 3 participants, neither the asynchrony nor the explicit response predicted the implicit response. For a further participant, the addition of the explicit response to the model rendered both the asynchrony variable and the explicit response variable non-significant, despite making a significant improvement to the model and leading to a significant model. Running two separate regressions with the asynchrony as the IV revealed that AV asynchrony significantly predicted both explicit and implicit responses. Although the correlation between the implicit and explicit responses was moderate (r=.338), it is possible that the variance explained by asynchrony is the same as the variance explained by the explicit variable, thus leaving both non-significant predictors.

### 2.4.3 SUMMARY

To summarize the above results, it seems that for McGurk integration, a rather large majority (70%) of individual data sets follow the pattern whereby implicit responses are independent from explicit responses, when absolute AV asynchrony is partialled out. In 19% of individuals, the explicit response predicted the implicit response over and above absolute asynchrony and in a further 7%, both the explicit response and absolute asynchrony predicted unique variance in the implicit response. In contrast, the explicit responses were independent from implicit responses in only 21% of the Stream-Bounce data sets. In 39% of the data sets, which make up the majority, the explicit response predicted the implicit response over and above absolute asynchrony and in a further 7% both the explicit response and absolute asynchrony predicted unique variance in the implicit response. Implicit data also tended to be independent from absolute asynchrony in more Stream-Bounce data sets (18%) compared to McGurk data sets (4%) suggesting that the effect of synchrony on AV integration in the stream bounce illusion is less robust than in the McGurk illusion. The Stream-Bounce dual-task is more prone to response bias, as the participant might assume that if two disks bounce off each other, the sound that would have resulted from the collision must have happened at the same time as the collision. As will be covered in detail in the discussion later, the McGurk paradigm tends to be less prone to such response bias and such bias can be spotted in the data.

To conclude, the majority of individual regressions in the McGurk data (70%) suggest that implicit and explicit responses tend to be independent of each other, although there are deviations from this in a proportion of individuals (26%) whose implicit responses are related to their explicit responses, when asynchrony is partialled out. The Stream-bounce data sets seem to be more unreliable and the predominant pattern of results - suggesting that the two responses are dependent of one another - comprises a smaller and less impressive proportion of total responses (39%). The latter result could however be due to the Stream-Bounce paradigm being rather prone to response bias and the illusion being less compelling.

# 2.5 DISCUSSION

Chapter 2 addressed whether AV information needs to be explicitly perceived as synchronous in order to be integrated by examining the relationship between estimates of ePSS and iPSS. Implicit PSS and ePSS measures were derived using a dual-task paradigm. A significant negative relationship between measures of ePSS derived from TOJs and measures of iPSS derived from both McGurk and the Stream-Bounce illusion was found. Estimates of ePSS derived from SJs and measures of iPSS were not significantly related, and the subject-by-subject analyses suggest that when response bias is ruled out, explicit timing judgements do not predict implicit judgements over and above asynchrony. These results do not support accounts which argue that explicitly perceiving synchrony in auditory and visual events leads to the assumption that the stimuli belong together, in turn leading to their integration into a unitary AV percept (Welch & Warren, 1980) and that AV speech information is more likely to be perceived as synchronous as a result of the unity effect (Vatakis, Ghazanfar, et al., 2008; Vatakis & Spence, 2008a). Under these accounts, the AV relative timing perceived most often to be synchronous would be expected to be the relative timing at which AV information is integrated most often. In other words, such accounts would predict a positive correlation between measures of iPSS and ePSS.

## 2.5.1 SJ EPSS AND IPSS

It is widely accepted that audiovisual integration depends on whether auditory and visual events are judged to belong together. According to this view, whilst an observer's assumption of unity can be promoted by AV synchrony, the assumption of unity can also affect whether AV events are perceived to have occurred at the same time (Spence, 2007; Vatakis, Ghazanfar, et al., 2008; Welch & Warren, 1980). Under this premise, the asynchrony which promotes maximum AV integration should be consistent with the AV asynchrony which maximally promotes the perception of AV synchrony, within individuals. In other words, estimates of iPSS and ePSS should correlate positively. This would be especially predicted for iPSS and ePSS measures derived from judgements on speech stimuli, as speech has been argued to afford a ‘special’ mode of perception and thus be more likely to promote the assumption of unity (Vatakis, Ghazanfar, et al., 2008). Instead, the null relationship observed between SJ ePSS and iPSS for both non-speech and speech AV stimuli, as well as the subject by-subject analyses carried out in section 2.4 indicate that AV information is not necessarily perceived as synchronous most often when it is also integrated most often, at least when the AV stimuli are not prone to response bias. This suggests that the underlying temporal mechanisms of AV integration and explicit synchrony judgements may not be shared. The subject-by subject analyses on the Stream-Bounce data may however suggest that when uncertain about their perception, individuals might consciously use one percept (e.g. bouncing) to make a decision about the other (e.g. synchrony).

#### 2.5.1.1 CAN RESPONSE BIAS ACCOUNT FOR THE NULL RELATIONSHIP BETWEEN SJ-EPSS AND IPSS?

One argument is that when judging whether a stimulus pair is synchronous, two criteria are used, one for visual leading and another for visual lagging stimuli (Yarrow et al., 2011). If the observer’s perception of AV relative timing for a given stimulus pair falls within these two criteria, then a ‘synchronous’ response will be made and if it falls outside the two criteria, then the stimulus will be judged as asynchronous. Alternatively, we may possess a directionless sense of what is synchronous (Vroomen & Keetels, 2010a), which might mean that we use one criterion for judging synchrony. Whether we use one or two criteria, inconsistency or random variability within and/or between individuals in these criteria could have resulted in less reliable estimates of ePSS in the SJ task, concealing a positive relationship between iPSS and ePSS estimates.

On the other hand, response bias such as the tendency to press a particular button or to switch between buttons is unlikely to account for the null relationship observed between iPSS and ePSS. Participants might be more likely to press the same button for both judgements, for instance by pressing ‘stream’ after they had pressed ‘synchronous’ or ‘bounce’ after they had pressed ‘asynchronous’. This would lead to inverse functions for the ‘bounce’ responses and would be easily noticed (and was checked for). Alternatively participants might be more likely to switch buttons, for instance pressing ‘synchronous’ followed by ‘bounce’ and asynchronous followed by ‘stream’. This would in principle strengthen a positive relationship between iPSS and ePSS rather than conceal it. The same would happen if responses were influenced by demand characteristics, whereby participants thought it was ‘correct’ to respond ‘bounce’ when the beep occurred at the same time as the point at which the disks crossed over.

## 2.5.2 TOJ ePSS AND iPSS

The asynchrony optimal for AV integration was negatively related to PSS measures derived from TOJ. In other words, participants who optimally integrate AV information when the visual stream leads the auditory were more likely to explicitly perceive AV synchrony when the visual stream *lags* the auditory, and vice versa. Additionally, the more auditory lag participants require to integrate AV information most often, the more auditory lead they need to explicitly perceive AV synchrony. Across implicit and explicit judgements, the same pair of AV stimuli was therefore perceived to be separated by two different AV asynchronies at the same, and these estimates seem to be repelled from one another. These findings point towards a dual perception in AV timing, similar to those reported by Soto-Faraco & Alsius (2007; 2009) and Martin et al. (2012). Previous findings of subjective disunity averaged across participants are however relatively small and could be accounted by response bias. As mentioned in Chapter 1 for example, PSS measures derived from TOJs are particularly prone to being affected by the observer’s decision criterion (Yarrow et al., 2011) and as a result, differences between this measure and the iPSS might be expected. The next section will discuss whether such biases would cause the negative correlation observed here between iPSS and TOJ ePSS.

### 2.5.2.1 CAN RESPONSE BIAS ACCOUNT FOR THE NEGATIVE RELATIONSHIP BETWEEN TOJ-ePSS AND iPSS?

Shifts in response criterion for the two concurrent judgements might account for differences observed between iPSS and ePSS measures in Soto-Faraco and Alsius

(2007). Similar explanations have been proposed to account for discrepancies observed between uncorrelated (van Eijk et al., 2008) measures of ePSS derived from TOJs and SJs (García-Pérez & Alcalá-Quintana, 2012; Schneider & Bavelier, 2003; van Eijk et al., 2008). Individual variability across participants in measures of iPSS and TOJ ePSS could be explained by prior entry effects (Spence & Parise, 2010; Spence, Shore, & Klein, 2001; Massimiliano Zampini et al., 2005). According to prior entry, attention speeds up perceptual processing, and results in an attended stimulus being perceived to have occurred earlier compared to an unattended stimulus presented at the same time. Thus, variability in the weighting of attention to one modality versus the other across participants might account for variability in PSS measures (Spence & Parise, 2010; Spence, Shore, & Klein, 2001). Whilst decisional or attentional biases could conceal a positive correlation between two measures, it is harder to explain how such biases could lead to the negative correlation observed between measures of iPSS and TOJ ePSS.

For example in the Stream-Bounce experiment, participants might have the tendency to press the same button for both responses, for instance, the ‘sound second’ and ‘bounce’ button. This would shift the psychometric function for TOJ horizontally, pushing ePSS values towards greater auditory lag but would only cause an increase in height of the Stream-Bounce function and not a horizontal shift which would repel the two measures. Alternatively, participants’ second key press might be dependent on their first key press. For example, they might be prone to either press the same key or to press the other key on the second question. This could increase or decrease the proportion of ‘bounce’ responses as a function of auditory lag, and shift the iPSS horizontally, but would have no effect on ePSS because the temporal order judgement is made first. Neither of these two button biases alone seem to account for the negative correlation observed between iPSS and TOJ ePSS. Adopting a combination of the two biases might in principle lead to a negative correlation, as the pressing the same button twice would shift ePSS, and a bias towards making a different key press for the second question would shift iPSS in the opposite direction. Maintaining such a convoluted response trend would however require effort on the part of the observer. Moreover, this combination of biases becomes even more complicated to implement in the McGurk experiment. In this context for example, the participant would have to know that a ‘da’ response represents AV integration in one condition and that it indicates that the visual information had no effect on the auditory in the other incongruent condition, otherwise their bias would be easily spotted from their data in the congruent AV syllable condition. This is highly unlikely as participants were naïve to the aim of the experiment. Furthermore, such perceptual discrimination would be extremely difficult, given the nature of the McGurk illusion. To conclude, it is unlikely that the negative correlation can be explained entirely by response bias, as adopting the right combination of biases would make a very intricate task for the observer.

Given that response and attentional biases cannot easily account for the negative correlation observed between estimates of AV synchrony derived from TOJ and AV integration, the following section will attempt to discuss the neural mechanisms that might give rise to multiple conflicting estimates of AV relative timing in reference to the same AV stream. The section will discuss the results in relation to existing theories of crossmodal temporal processing and conclude that these cannot account for the disunity in subjective AV timing and the negative correlation observed iPSS and TOJ-ePSS estimates observed here. The section will finally discuss a novel proposition that explains these results.

## 2.5.3 THEORETICAL ACCOUNTS

### 2.5.3.1 HOW MIGHT MULTIPLE ESTIMATES OF AV RELATIVE TIMING ARISE?

Some accounts of time processing describe specialised timing mechanisms which reside in cerebellum or basal ganglia (Ivry & Spencer, 2004), providing a common time code for multisensory events. Accounts in favour of such internal clocks or pacemakers however imply that their functions are stable and not necessarily adaptable to specific task demands (Nenadic et al., 2003). Another implication from such accounts is that a single central timing mechanism is responsible for estimating the relative timing of AV information. Under such accounts, it is difficult to envisage how several estimates of AV relative timing could arise for a single AV stream. The disunity in subjective timing across perception of temporal order and AV integration observed here does not by any means invalidate such accounts, and might arise if common time codes were used for reference, by other localised temporal processing mechanisms which processed temporal information specifically for the demands of the tasks they are dedicated to. Therefore, rather than constructing estimates of the relative timing of crossmodal events, it is possible that mechanisms supported by subcortical structures such as the basal ganglia and cerebellum provide information regarding time to other, localised, task-specific temporal processes.

Functional imaging studies have demonstrated that multisensory processes are associated with the activation of a distributed network of cortical structures, depending on the task being performed and type of stimuli being processed (Callan et al., 2003; Calvert, Hansen, Iversen, & Brammer, 2001; Ghazanfar & Schroeder, 2006; Jones & Callan, 2003; Sekiyama, 2003; Stevenson, van DerKlok, Pisoni, & James, 2011; Wright, Pelphrey, Allison, McKeown, & McCarthy, 2003). Different types of AV integration have also been shown to correlate with activity in distinct neural networks (Bertini et al., 2010). Temporal processing has been shown to engage a wide network of cortical and subcortical structures (Rubia & Smith, 2004) and is thought to be processed via an assembly of multiple overlapping mechanisms, which are stimuli and task-dependent (Eagleman, 2008; Wiener, Matell, & Coslett, 2011). It is therefore possible that the distributed activations observed in fMRI research might reflect multiple, task-specific temporal processing mechanisms.

If temporal processing were supported by multiple mechanisms, the question of how disunity in subjective synchrony could arise in the first place still remains. When visual and auditory events occur synchronously at source, the resulting information travels at different speeds externally and internally and as a result, reaches separate unisensory destinations in the brain at different times (King, 2005). Auditory and visual information then has to converge at crossmodal mechanisms to be processed as a unified percept. The location of the neural substrates which support these mechanisms will vary and depend on the combination of the processes required to complete the task being performed. These will be subject to different neural delays and as a result, these mechanisms are likely to produce various conflicting estimates of relative timing in reference to the same pair of AV stimuli, if AV timing is indeed processed by localised task specific mechanisms.

Under this account, it can be seen how the neural mechanisms underlying temporal order and AV integration might produce different estimates of AV synchrony. Inter-individual variability in these measures might be introduced by functional and structural brain differences across individuals, which might affect the relative speed of visual and auditory propagation, transduction and processing. To summarise, disunity of subjective AV timing across temporal order perception and AV integration might arise if different multisensory processes were each subserved by task-dependent temporally sensitive mechanisms, each receiving different information about the relative timing of the same audiovisual stream.

The above explanation alone does not account for a negative relationship observed between measures of iPSS and TOJ ePSS. The next section will discuss existing theoretical accounts of how the nervous system might account for internal and external AV delays in order to maintain temporal coherence. The section will conclude that these accounts alone cannot account for the negative correlation observed between TOJ-ePSS and iPSS. The section will propose that the antagonistic relationship might be indicative of renormalisation of multiple estimates of AV relative timing, relative to the mean, or in other words that AV relative timing within localised, task-specific temporal processing mechanisms is assessed in relation to the average relative timing across all active temporal mechanisms.

### 2.5.3.2 HOW MIGHT TEMPORAL COHERENCE BE ACHIEVED DESPITE MULTIPLE ESTIMATES OF AV RELATIVE TIMING?

In order to maintain temporal coherence despite multiple estimates of AV synchrony, adjustments to AV temporal estimates might be made in order to account for delays caused by external factors such as observer-stimulus distance (Harris, Harrar, Jaekl, & Kopinska, 2008) or to fit in with prior knowledge of the likelihood that events are synchronous or not (Miyazaki, Yamamoto, Uchida, & Kitazawa, 2006; Yamamoto, Miyazaki, Iwano, & Kitazawa, 2012). To account for internal factors, such as propagation times of sensory signals, delaying or speeding up one modality in relation to another (Luca et al., 2009; Navarra, Hartcher-O'Brien, Piazza, & Spence, 2009) or altogether recalibrating temporal codes (Fujisaki et al., 2004; Hanson et al., 2008) would result in a frequently occurring neural asynchrony to be perceived as synchronous. Recalibration mechanisms have been suggested by some to be supramodal and to maintain temporal coherence despite sensory lags via adjustment, or recalibration of sensory timing regardless of the modality to which sensory inputs belong to (Hanson et al., 2008). The notion of a central recalibration mechanism is rather difficult to fit with these results. Firstly, one implication of a supramodal recalibration mechanism such as the one described by Hanson et al. seems to be that having reached the nervous system, the relative timing of crossmodal signals would be assessed and the resulting estimate would either accepted if it fits in with the observer's prior knowledge of the world, or recalibrated if it does not. Under this account a pair of audiovisual streams would have no more than one estimate associated with it regarding AV relative timing, which is at odds with the observation that participants in the current study experienced a dual perception regarding AV relative timing. Under the notion of central recalibration of timing, iPSS and ePSS measures should agree as any discrepant estimates would be recalibrated towards one another. The renormalisation account however does not necessarily invalidate recalibration accounts as the latter might resynchronise estimates of AV timing within multisensory perceptual processes, rather than across.

The negative relationship observed between measures of TOJ ePSS and iPSS is difficult to account for under assumptions of unity and automaticity too, as these accounts would more likely predict a positive correlation between these measures. These results not only suggest that the process of temporal order and that of integrating AV information are likely to be supported by distinct synchronisation mechanisms, but that estimates derived from these mechanisms do not seem to be adjusted in a way which minimises discrepancies between them. As mentioned earlier, the proposal that subjective synchrony is achieved through multiple local, task-specific temporal mechanisms, each subject to different neural delays, alone does not explain the negative correlation. Whilst this explanation would account for differences between estimates of subjective AV relative timing, it does not account for the nature of the relationship observed and therefore needs a further assumption in order to do so, which will be discussed next.

#### 2.5.3.3 RENORMALISATION OF SUBJECTIVE AUDIOVISUAL TIMING

Patient PH, who was described in the introduction to this chapter, needed the visual information to lead the auditory stream by approximately 200ms in order to perceive the streams as synchronous, but the visual stream to *lag* the auditory by the same amount of time in order to integrate the two most often (Freeman et al., 2013). Across these two processes, PH’s average estimate of AV relative timing seems to be veridical, as the mean of the individual estimates is approximately zero, which represents AV synchrony. This also seems to somewhat apply to individuals tested here. The two mechanisms underlying the two tasks measured here are likely to be part of a larger network of localised mechanisms, which might together produce a distribution of multiple estimates of AV relative timing. It is also likely that for each individual timing mechanism contributing to this distribution, there is uncertainty regarding the objective timing of the events being processed. Without knowing which estimate of relative timing is the most veridical or the timing at which auditory and visual events occurred outside the nervous system, the most reliable reference, or in other words the neural asynchrony which most likely relates to objective synchrony, would be the average of all available estimates of relative timing. Rather than being adjusted towards the average estimate of AV relative timing, renormalisation proposes that subjective estimates of the relative timing between sensory events are perceived *relative* to the overall average (see Figure 2.26 on page 90). In PH’s case, a disruption in either AV integration or temporal order processing might lead to an extreme measure of iPSS or ePSS, respectively. It is possible that the arrival time of auditory information at AV integration mechanisms was slowed, leading to an extreme visual leading subjective estimate of AV synchrony within AV integration mechanisms, because PH’s lesion was located in subcortical structures previously implicated in auditory processing. As shown in Figure 2-26, the addition of this extreme visual-leading AV synchrony estimate would widen the overall distribution of AV asynchronies produced by the different localised timing mechanisms, as well as shift the average asynchrony estimate towards the extreme estimate. If each local estimate was assessed relative to this mean, then localised estimates which were assessed to be visually lagging to begin with would now be assessed as more visually lagging than they were before, relative to the new average estimate of AV synchrony. In other words, if audition was slowed and auditory signals arrive particularly late at one mechanism, the overall spread of the distribution is altered and auditory signals are then seen as arriving particularly early in other mechanisms. The same would apply in situations where the visual information is consistently slow for a given process; visual information from other mechanisms on the other side of the distribution (i.e. ‘auditory arrives second’ side of bottom right of Figure 2.26) would then seem to arrive particularly fast.

At first glance, the account of renormalisation (Freeman et al., 2013 illustrated in Figure 2.26, on the next page) does not seem to neatly fit in with accounts of recalibration, adaptation or temporal ventriloquism. However, it must be stressed that such accounts have primarily stemmed from studies that adapt and test within single multisensory processes. Renormalisation aims to account for how AV timing is assessed across *different* multisensory processes which are likely to compute different estimates of subjective synchrony. Renormalisation allows for the possibility that sensory timing *within* specific mechanisms, such as those dedicated to processing synchrony (Vatakis, Navarra, et al., 2008) or temporal order (Vatakis et al., 2007), and even AV integration (Asakawa, 2008), can be recalibrated or adapted, but also describes how sensory timing within a particular mechanism would be assessed in relation to the estimates of other mechanisms, given a

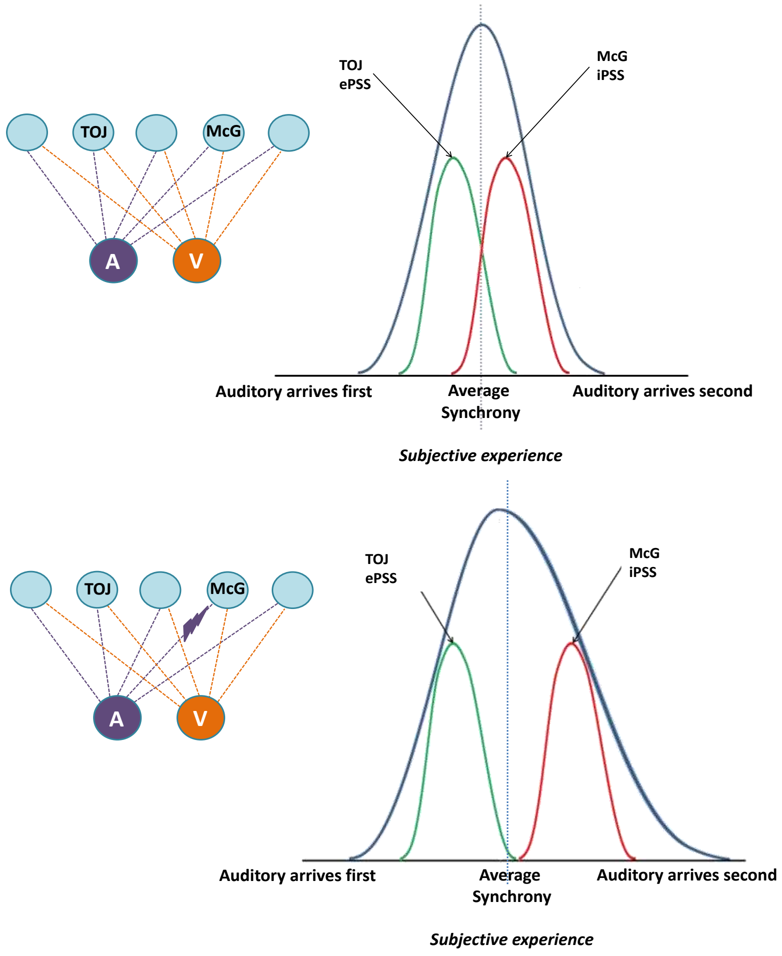


*Figure 2-26: **Temporal renormalisation theory**. Hypothetical relationship between neural and subjective audiovisual asynchrony. Top left: signals from synchronous auditory and visual stimuli (represented by purple and orange disks) converge on different audiovisual mechanisms in the brain via different routes (blue disks). For individual mechanisms the actual stimulus timing cannot be dissociated from the propagation latency. Top right: schematic of the evoked distribution of neural asynchronies, across mechanisms, plotting probability of different asynchronies, as a function of neural asynchrony, with increasing delays of auditory signals relative to visual towards the right. Within this distribution, evoked distributions within TOJ and McGurk mechanisms are shown by the green and red curves, of which the peaks represent the subjective synchrony estimates, ePSS and iPSS, respectively. The x-axis refers to the subjective experience of auditory led, simultaneity, or auditory lag, given these different neural asynchronies. The neural asynchrony at the central tendency of the distribution is the one which relates most reliably to the objective timing of the auditory and visual stimuli, after delays within individual mechanisms have been averaged out. Following experience with this distribution in natural contexts where objective synchrony is likely, tasks probing mechanisms registering asynchronies near this average may evoke perception of synchrony (marked with a dotted line and 'average synchrony'); asynchronies registered within other mechanisms are perceived in proportion to their distance from the average. Lower left: an example where auditory inputs to a subset of mechanisms (towards the right) are particularly delayed. For patient PH it is assumed that these mechanisms contribute to the temporal tuning of the McGurk illusion (labelled McG; see main text), while mechanisms involved in TOJ are preserved. Lower right: the distribution resulting from delayed auditory input for the McGurk task. The mean of the distribution has shifted towards the auditory-lagged mechanisms serving the McGurk task (labelled McG). The perceived asynchrony within each mechanism is renormalized to this new distribution mean. The result is that neural asynchronies for unaffected mechanisms (here labelled TOJ) are now perceived as more auditory-leading.*

disruption or recalibration of timing in other mechanisms. It would for example predict that if one given mechanism were exclusively adapted to a given AV asynchrony, this adaptation might result in a shift in its estimate of subjective synchrony towards the adapted asynchrony, and a shift away from it for other, unadapted mechanisms, following renormalisation.

### 2.5.4 CONCLUSION

To conclude, it seems that the nature of the relationship between explicit and implicit iPSS and ePSS derived concurrently is dependent upon the type of explicit temporal judgement task that is being performed, but not on the type of stimuli employed. When performing temporal order judgements and integrating AV information concurrently, individuals seem to experience disunity in the relative timing of sensory information. Estimates of subjective synchrony derived from temporal order judgements and AV integration correlate negatively, suggesting that AV synchronisation underlying these processes is supported by independent mechanisms, whose estimates of AV relative timing are renormalized relative to the average timing across all other synchronisation mechanisms. In contrast, a null correlation was found between estimates of ePSS derived from SJs, and iPSS. The failure to find a positive correlation is at odds with accounts of unity which argue for a mutual dependence between explicit subjective synchrony and AV integration, and suggest that AV temporal processing is supported by multiple, task-specific mechanisms.

# 3 CHAPTER 3: SIZE DOES MATTER: MORPHOLOGICAL CORRELATES OF IMPLICIT AND EXPLICIT TEMPORAL PROCESSING AND AUDIOVISUAL INTEGRATION

## 3.1 INTRODUCTION

In Chapter 2 estimates of implicit subjective synchrony of audiovisual (AV) speech were found to correlate negatively with estimates of explicit subjective synchrony derived from temporal order judgements (TOJ) and were not statistically related to those derived from synchrony judgements (SJ). This observation indicates that AV synchronisation across different AV speech processes might be achieved via multiple, rather than a single, common temporal mechanism. Furthermore, the failure to find a positive relationship between implicit and explicit measures of subjective synchrony in the previous chapter suggests that these mechanisms are likely to be subject to different neural delays. Given this evidence, the current investigation is the first to explore whether performance in AV temporal discrimination across implicit and explicit tasks is statistically dependent or independent and critically, whether it is related to individual differences in the structure of the same or distinct anatomical brain areas. A similar aim of the chapter was to examine the relationship between brain structure and individual susceptibility to the McGurk illusion.

Accounts of unity argue that integration of multiple crossmodal streams is contingent upon on the perceiver’s assumption that they belong together. Whilst this assumption can be promoted by AV synchrony, it in turn can affect the perception of whether two or more events occurred at the same time, if they have been integrated into a unitary percept (Spence, 2007; Vatakis, Ghazanfar, et al., 2008; Welch & Warren, 1980). Under this premise, AV integration should be optimal when AV streams are perceived to be synchronous and therefore individual ability to discriminate between synchronous and asynchronous AV stimuli should be consistent across AV integration and explicit timing judgements, in particular when the two are being performed at the same time. Behavioural measures of the ability to discriminate between AV synchrony and asynchrony derived from McGurk judgements should therefore correlate positively with measures derived from explicit timing tasks. Furthermore these measures should in turn be related to structural variability within the same brain areas.

Alternatively, AV synchronisation across different processes might be supported by distinct underlying temporal mechanisms, as suggested by the results of Chapter 2. If this were the case, these mechanisms may not be subject to the same factors underlying temporal processing performance, and temporal discrimination measures derived from McGurk functions would be uncorrelated to those derived from temporal order and synchrony judgements. This would mean that the ability to explicitly discriminate synchronous from asynchronous AV information may not be dependent on, or be positively related to the degree to which synchronous and asynchronous AV information is discriminated during AV integration, or vice versa. If this were the case, sensitivity in discriminating synchronous from asynchronous AV information in the McGurk task would be statistically independent from the sensitivity with which synchrony is discriminated from asynchrony during explicit temporal judgements. Furthermore, if individual differences in temporal discrimination across these tasks are related to structural variability of distinct anatomical areas, it would indicate that AV temporal processing across implicit and explicit timing tasks is supported by distinct underlying temporal mechanisms.

The following introduction will first discuss functional imaging research which has attempted to map the underlying mechanisms of AV synchrony processing and/or perception to the human brain. The review will then evaluate functional imaging research on the processes underlying the detection of AV temporal order. Lastly, the introduction will outline and discuss literature on the functional correlates of the McGurk effect. The review of functional imaging studies investigating AV integration is restricted to studies that utilised McGurk effect to induce or measure AV integration. Functional MRI literature on AV timing is rather limited and thus the review will include studies that have utilised a range of AV stimuli, including AV speech and non-speech stimuli and unisensory stimuli.

### 3.1.1 NEURAL CORRELATES OF AUDIOVISUAL SYNCHRONY PROCESSING AND PERCEPTION

Studies concerned with functional mapping of AV synchrony processes typically compare neural activity which correlates with perceived or physical AV synchrony to activity which correlates with perceived or physical AV asynchrony, respectively. Other paradigms compare brain activity measured during trials in which participants perform synchrony judgements to brain activity measured during trials in which participants perform a task that does not require temporal processing. In one of the most comprehensive studies of AV synchrony perception to date, Noesselt, Bergmann, Heinze, Münte, & Spence (2012) investigated the neural correlates of processes underlying the perception and processing of AV (a)synchrony and AV temporal order in continuous AV speech. The authors measured BOLD responses during conditions in which AV speech was presented synchronously, as well as in which the auditory led and lagged the visual component of the stimulus. Inside the scanner, participants reported when their perception regarding the synchrony of the stimulus changed, for example from synchronous to asynchronous or vice versa. By doing so, the authors were able to measure BOLD activation that corresponded with physical as well as to perceived AV (a)synchrony. The authors reported that changes in the temporal properties of AV stimuli resulted in increased activity in sub-regions of the posterior-middle STS bilaterally, irrespective of the observer’s subjective perception of relative timing. These activations however were further enhanced if the observer’s perception was congruent with the physical temporal properties of the stimuli. For example, activation to auditory-leading asynchronies was higher if the participants concurrently perceived the stimulus pair as auditory-leading compared to if they had judged its temporal order incorrectly. Furthermore, exposure to the different types of stimulus pairs resulted in increased activity in distinct subregions. In other words, processing auditory-leading, synchronous and auditory-lagging AV information correlated with activity in distinct regions of the STS, each selective to a particular type of AV relative timing. The perception of asynchrony in general was also related to enhanced BOLD activity in prefrontal regions as well as stronger functional connectivity between the STS and these areas. In summary, this study indicates that the relative timing of AV information is processed via a temporal-prefrontal network consisting of multiple neuronal populations, each responsive to different temporal characteristics and that connectivity within this network and its dynamics are modulated by the observer’s perceptual state.

The superior temporal cortex, superior colliculus and cerebellum (Stevenson, Altieri, Kim, Pisoni, & James, 2010) have also been implicated in the perception of synchrony in AV speech. Stevenson et al. (2010) compared BOLD responses measured during presentation of synchronous and asynchronous audiovisual speech clips consisting of single, familiar words on which subjects performed semantic categorisation judgements. Unlike Noesselt and colleagues, Stevenson et al. did not measure perceived synchrony inside of the scanner, but did report that on average, the perception of the stimulus measured prior to scanning reflected its physical temporal properties. Within the superior temporal cortex bilaterally, two areas showed differential patterns of activation to synchronous versus asynchronous AV stimuli. One of these areas only showed greater activation to synchronous AV stimuli, and another showed greater BOLD activation as the temporal offset between the auditory and visual components of the stimuli increased. These results concur with those of Noesselt et al. (2012), who also reported separate areas for processing AV synchrony and asynchrony. Stevenson et al. cautiously proposed that the BOLD activation corresponding to the increase in AV asynchrony could reflect one of two things: either an area that responds to multiple incoming sensory streams that have not been integrated, or which reflects increasing demands imposed on the process of AV integration by the increase of AV asynchrony. However in light of Noesselt et al.'s results, it is possible that processing the temporal characteristics of AV stimuli recruits distinct neuronal populations depending on whether the streams are synchronous or asynchronous. Along with Noesselt et al.'s findings, these results lend some support for the notion proposed by Roach, Heron, Whitaker, & McGraw (2011), that AV relative timing is estimated from the distribution of activity derived from populations of neurons, each responsive to different AV temporal delays. Other areas reported by Stevenson et al. to show increased activity as a result of AV synchrony following a whole brain analysis included bilateral superior colliculus, posterior Fusiform gyrus, lateral occipital complex and extrastriate visual cortex.

One can infer at least in part, that the subregions within the STS, prefrontal regions and insula identified by Noesselt et al.(2012) as showing patterns of activations related to perceptual states, are likely to reflect cortical areas that support mechanisms underlying explicit perception of synchrony. This is because participants in this study were required to make explicit synchrony judgements during scanning and the resulting activations were analysed as a function of the observer's perceptual state, and not simply of the physical properties of the stimuli. Conversely, Stevenson et al.'s participants were not required to attend to the temporal properties of the AV stimuli, but instead performed a semantic categorisation task. It is thus unclear as to whether the relative timing or synchrony of the stimuli was being processed implicitly and/or explicitly by participants in this study. It may be tempting to infer that the (a)synchrony of the stimuli was being processed implicitly because participants were not overtly asked to attend to it, but one cannot confidently assume what participants were attending to in the scanner, as no measurement was taken to confirm it. As timing judgements were performed outside of the scanner prior to the main experiment, some participants may have attended to the timing of the stimuli having had to do so in the behavioural experiment, but it is just as likely that they did not. It is therefore ambiguous as to whether the areas identified by Stevenson which were not also identified by Noesselt et al. reflect areas related to implicit or to explicit temporal processing.

Stevenson et al.’s analyses were also performed as a function of the physical (a)synchrony of the stimulus and not of perceptual states. This is problematic as perceived synchrony of AV information can be unstable (Kanai, Sheth, Verstraten, & Shimojo, 2007), therefore one cannot assume confidently that a pair of stimuli which were on average perceived as synchronous outside the scanner was always be perceived to be so inside the scanner.

Contrasting BOLD activity between auditory-leading and lagging conditions is a good method to reveal activations which are likely to reflect activity of distinct neuronal populations responsive to different AV neural delays. However, findings based on contrasts between activity correlated with presentation of synchronous and asynchronous AV stimuli are likely to be confounded by processes underlying AV integration. This consequently means that at least some of the areas reported to show higher levels of activity during synchronous presentations of AV stimuli by both Noesselt et al. (2012) and Stevenson et al. (2010) may actually represent areas that support AV integration as opposed to synchrony processing, as the visual and auditory components of the AV stimuli used in these studies were congruent and therefore likely to be readily integrated.

Bushara, Grafman and Hallet (2001) reported the inferior frontal gyrus, cerebellum and inferior parietal lobule to be involved in AV synchrony processing of non-speech stimuli. They measured brain activity using PET whilst participants performed synchrony judgements on pairs of simple auditory and visual stimuli. AV stimuli presented at 6 different asynchronies were blocked according to whether the visual component led or lagged the auditory stimulus and were interleaved with pairs of synchronous AV stimuli. PET activity measured during the synchrony judgement condition was split between trials in which the AV stimuli were synchronous and trials in which the stimuli were asynchronous. Only the activity from the latter used in the analysis; activity was averaged across these trials and compared against activity measured during trials in which participants judged the colour of the visual component of a synchronous AV stimuli pair. According to the authors, the control task enabled the exclusion of activity related to sensorimotor responses and attention to AV stimuli from activations measured in the experimental condition. The authors concluded that AV synchrony detection employs a large cortical network comprising insular, posterior parietal, prefrontal, and cerebellar areas. Within this network, activity in the right insula was reported to be positively related to task difficulty.

Bushara et al.’s (2001) choice of task for the control condition however calls for some caution in the interpretation of their results. In the experimental condition, participants were required to make a judgement based on both components of the AV stimuli and therefore had to attend to stimuli of two modalities. Conversely, in the control condition they were only required to make judgements on the visual stimulus meaning that they did not need to attend to both the visual and auditory components of the stimulus. Due to the inconsistency in attentional demands across the control and experimental conditions, it is unclear as to whether some of the activations reported were actually related to detecting AV (a)synchrony, or whether they were in fact related to an increase in attentional demands in the experimental condition and consequently unrelated to temporal processing. Furthermore, brain activity in the baseline condition was measured from trials in which AV stimuli were always synchronous, and therefore likely to be integrated into and perceived as one multisensory stream. Conversely, the activity in the synchrony detection condition was measured from trials in which the AV stimuli were asynchronous, and thus less likely to be perceived as a unitary AV percept. This means that some of the increased activity in the synchrony detection condition relative to the control could be related to a difference in processing multiple streams of unisensory information versus a single multisensory stream, a process which may not strictly be associated with AV temporal processing. Since its publication however, other studies have associated areas similar to those reported by Bushara et al., to be related to AV synchrony processing, namely the Insula and cerebellum (Stevenson et al., 2010). In addition, one study which used only visual stimuli to measure activation during synchrony judgements reported activation of the middle and superior temporal gyrus (Lux, Marshall, Ritzl, Zilles, & Fink, 2003). A full list of brain areas reported to be associated with AV and unisensory synchrony perception and/or processing can be found below in Table 3-1.

| Anatomical structure | L/R | Contrast / task | Reference |
|---|---|---|---|
| **Superior temporal sulcus** | Bi | Percept: Synch > asynch | (Noesselt et al., 2012) |
| | Bi | Percept: Asynch >synch | (Noesselt et al., 2012) |
| | R | Physical: synch>asynch | (Noesselt et al., 2012) |
| **Superior temporal cortex** | R | Synch/asynch detection AV speech | (Stevenson et al., 2010) |
| **Superior temporal gyrus** | L | synch V asynch detection > orientation detection | (Lux et al., 2003) |
| **Insula** | Bi | Synch/asynch detection > colour detection of synch stimuli (C) | (Bushara, Grafman, & Hallett, 2001) |
| **Inferior frontal gyrus** | R | Synch/asynch detection > colour detection of synch stimuli (C) | (Bushara et al., 2001) |
| | L | Synch/asynch detection > colour detection of synch stimuli (C) | (Bushara et al., 2001) |
| | L | synch V asynch detection > orientation detection | (Lux et al., 2003) |
| **Middle Frontal gyrus** | L | synch V asynch detection > orientation detection | (Lux et al., 2003) |
| **Prefrontal Cortex** | Bi | Percept: Asynchrony > synchrony | (Noesselt et al., 2012) |
| | R | Physical: synch> asynch | |
| **Insula** | R | Percept Anterior: Asynchrony > synchrony | (Noesselt et al., 2012) |
| **Inferior parietal lobule** | R | Synch/asynch detection > colour detection of synch stimuli (C) | (Bushara et al., 2001) |
| **Cerebellum** | L | Synch/asynch detection > colour detection of synch stimuli (C) | (Bushara et al., 2001) |
| | | Synch/asynch detection AV speech | (Stevenson et al., 2010) |
| **Superior colliculus** | Bi | Synch/asynch detection AV speech | (Stevenson et al., 2010) |
| **Temporo-parietal Junction** | L | synch V asynch detection > orientation detection | (Lux et al., 2003) |
| **Parietal insular cortex** | R | Visual synch V asynch detection > orientation detection | (Lux et al., 2003) |
| **Lateral Occipital complex** | Bi | Synch/asynch detection AV speech | (Stevenson et al., 2010) |
| **Extrastriate cortex** | Bi | Synch/asynch detection AV speech | (Stevenson et al., 2010) |

*Table 3-1: Areas identified in fMRI studies: AV Synchrony/asynchrony detection*

A recurring issue in AV synchrony research which was mentioned previously and is worth elaborating is that AV synchrony often results in AV integration. As a result, it is likely that correlating brain activity with conditions in which AV stimuli are presented synchronously is not only likely to identify areas related to synchrony processing or perception, but also likely to reveal activity that is related to the integration of AV information. This would not be problematic if synchronisation and integration of AV information was supported by the same underlying mechanisms and cortical structures. However, there is some evidence that this may not be the case and that these two complementary AV processes are supported by distinct neural substrates.

Stevenson, vanDerKlok, Pisoni, & James (2011) set out to investigate the possibility that the integration and the synchronisation of AV information are supported by distinct neural structures. The authors identified two subregions within the Superior Temporal Cortex (STC), each selective for either AV synchrony or integration. Behavioural psychophysics were used prior to scanning, to measure whether participants perceived AV information as a unified AV percept or as two distinct events, as a function of AV asynchrony. The stimuli consisted of AV clips in which a female uttered a monosyllabic word, and each was presented at various SOAs ranging from 300ms auditory-leading to synchronous. The auditory-led asynchrony at which the AV stimulus was equally likely to be perceived as a unified percept and as two distinct events was then read from the resulting cumulative Gaussian curve. This then provided an ambiguous AV asynchrony for each individual at which AV stimuli could be perceived as a unified percept on some trials and as two separate events on others. Inside the scanner, the same AV stimuli were presented at this ambiguous asynchrony, as well as with the audio stream leading the visual by 400ms or occurring synchronously. This allowed comparisons between brain activity during trials in which the information was asynchronous and integrated versus brain activity during trials in which the information asynchronous and not integrated, keeping all other stimulus properties constant, including objective AV asynchrony, for this particular analysis. The 400ms auditory-led stimuli allowed for comparisons between activity to synchronous and unambiguously asynchronous AV stimuli.

Participants performed the same judgements as those carried out in the session prior to scanning, in which they indicated whether they perceived each stimulus as one or two events. Stevenson et al. (2011) then searched for increases in brain activity which correlated with objective synchrony and with integration of AV information, in separate analyses. One subregion of the STS was reported to show increased activity for synchronous compared to unambiguously asynchronous stimuli, but not to show the same differential activity when contrasts were made between trials in which the AV stimuli were integrated and trials in which they were perceived as two events. This region was consequently labelled the ‘Synchrony Defined Multisensory STC’ (shown in Figure 3.1, bottom). A different subregion showed increased activity during trials in which the stimuli were perceptually unified compared to trials in which they were not, but did not show any preferential activation for synchronous over asynchronous stimuli, and was consequently labelled the ‘Bimodal Multisensory STC’ (shown in Figure 3.1, top). This area also did not show any differential activity when integrated synchronous trials were compared to ‘integrated-asynchronous’ trials. This study therefore indicates that although the STC is implicated in both the synchronisation and integration of AV information, these complementary AV processes are likely to be supported by distinct subregions in this anatomical area and emphasizes the importance of distinguishing brain activity related to synchronisation of AV information from that related to its integration. The study

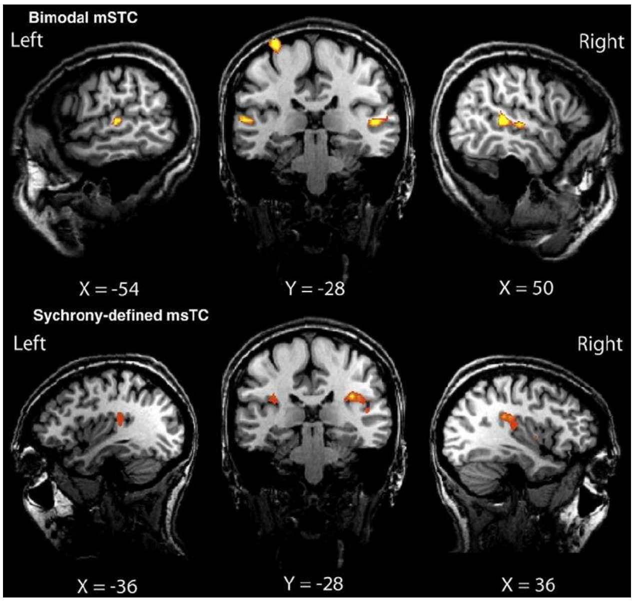


*Figure 3-1: Figure borrowed from Stevenson et al. (2011) showing (top) the areas with greater activity for integrated versus non-integrated AV stimuli and (bottom) areas with greater activity for synchronous versus asynchronous AV stimuli*

does not however address the current question of whether implicit temporal processing for AV integration and explicit synchronisation are dependent on common or distinct mechanisms.

### 3.1.2 NEURAL CORRELATES OF TEMPORAL ORDER

Noesselt et al.'s (2012) fMRI study on the neural correlates of temporal processing described previously also examined BOLD responses as a function of the temporal order in which AV speech stimuli were presented. To recap briefly, in one of the contrasts Noesselt et al. compared average BOLD responses measured during trials in which visual speech information either led or lagged the auditory stream to responses measured during trials where the AV streams were presented synchronously. These contrasts revealed that different areas showed increased activity as a function of whether the visual stream lagged, led or was presented synchronously with the auditory stimulus. Areas showing greater activation in the visual leading trials compared to synchronous ones comprised the right anterior and bilateral middle STS, bilateral prefrontal cortex and bilateral anterior insula. During visually lagging trials, right anterior STS, right posterior/middle STS, left-middle and left-posterior STS, right Precentral gyrus, left anterior insula and left prefrontal cortex showed enhanced activity compared to synchronous AV trials.

Noesselt et al.’s study provides an insight into the areas that may be of interest for the current structural correlation with performance in explicit synchrony judgements as well as for specific predictions about which areas may be related to performance in TOJs, as participants were asked to report their subjective perception of the stimuli. The perception and processing of AV synchrony were associated with activity in the posterior-middle STS. Activity related to the processing and perception of temporal order, that is, auditory-leading and lagging stimuli, was associated with increased activity in prefrontal areas as well as subregions of the STS, different to those whose activity was associated with the perception of synchrony.

Adhikari, Goshorn, Lamichhane, & Dhamala (2013) also investigated the neural correlates of AV temporal order perception. Outside the scanner, participants performed 3-alternate forced-choice temporal order and synchrony judgements on beeps and light flashes of varying AV asynchrony. Inside the scanner, participants were directed to perceive auditory-leading stimuli as either synchronous or visually-leading. Asynchrony conditions were blocked in terms of absolute AV asynchrony. That is, in one block beeps and flashes were presented synchronously or separated by 100ms in each direction (visually-leading or auditory-leading) and in another block the absolute asynchrony was 200ms. Average BOLD activity during trials in which participants were instructed to perceive auditory-leading AV stimuli as asynchronous and to perceive them as synchronous was contrasted. Activations related to perceived auditory-leading AV asynchrony consisted of clusters in temporal parietal junction and frontal areas. Clusters within these areas included the right superior temporal gyrus, inferior parietal lobe and supramarginal gyrus, left medial frontal gyrus, dorsolateral prefrontal cortex inclusive of the right middle frontal gyrus, and left intraparietal lobule.

Adhikari et al.'s (2013) methodology raises some concern regarding validity. Firstly, it is unclear why participants were instructed to perceive auditory-leading stimuli as synchronous or auditory-leading, instead of being asked to report their perceptual experience. *Trying* to perceive temporal order in a prescribed way and actually perceiving it are two qualitatively different tasks, likely to recruit different neural processes. Secondly, it is unclear why the only contrast carried out was between auditory-leading and synchronous conditions. Note that Noesselt et al. (2012) reported distinct activation patterns for auditory-leading, lagging and synchronous stimuli. It is therefore likely that, even if the task instructions were a valid way of measuring brain activity correlated with temporal order perception, the activation map revealed by Adhikari et al. (2013) to be associated with temporal order processing and perception would be incomplete due to the lack of contrasts performed for auditory-lagging AV stimuli.

The functional correlates of temporal order processing have also been investigated in paradigms employing unisensory stimuli, in which explicit temporal judgements were made within modalities, either on visual-only or auditory-only stimuli. The areas reported to show increased activity for both visual temporal order and auditory temporal order are the left supramarginal gyrus and inferior frontal lobe. Moser, Baker, Sanchez, Rorden, and Fridriksson (2009) compared BOLD activity between trials in which participants performed TOJs on auditory syllables and trials during which they identified the gender of the speaker. The authors controlled for the lack of phonological processing in the gender ID condition and its presence in the TOJ condition using an additional syllable identification condition. Contrasts between activity during the TOJ task and the control conditions revealed that auditory TOJs recruit the left IPL, specifically the Supramarginal gyrus, as well as the inferior/posterior frontal lobe, bilaterally.

Using visual-only stimuli, Davis, Christie and Rorden (2009) presented participants with two temporally offset visual stimuli presented in different locations on which they were required to perform either a shape discrimination task or a TOJ. Activation measured during the shape discrimination task was compared against activity during the TOJ. In addition to some of the areas reported by Moser et al. (2009) (see Table 3-2 and 3-3 on the next page for full list), Davis et al. reported that the inferior frontal lobe, superior frontal gyrus, intraparietal sulcus, temporoparietal junction and thalamus showed increased activity during the temporal order condition, relative to shape discrimination.

The perception and processing of AV synchrony and temporal order seems to engage various anatomical structures, most of which are located in temporal, frontal and parietal cortical areas (see Figures 3-2, 3-3 and 3-4 starting on page 107 for summary). The specific networks identified to be involved in these processes however vary quite substantially from one study to another. Furthermore, although the networks identified are broadly consistent in terms of the cortical lobes within which their components are located, no two studies seem to have reported the same specific structure. This could presumably be a result of differences in the type of paradigms, stimuli and tasks that are employed across the literature. There are some areas, namely the STS, STG, IPL, MTG, IFG, the temporo-parietal junction, the insula and prefrontal cortex which have been implicated in both the processing of AV synchrony and AV temporal order.

| Anatomical structure | L/R | Task/Contrasts | Stimuli | Reference |
|---|---|---|---|---|
| **Superior temporal sulcus** | R | Perceptual state:<br>Auditory lead vs synch vs visual lead | AV Speech | Noesselt et al., 2012 |
| | R | Anterior-posterior-middle:<br>Physical Auditory lead | | |
| | L | Posterior<br>Physical Auditory lead | | |
| | L | Middle STS<br>Physical Auditory lead | | |
| | Bi | Posterior<br>Percept: visual lead > synch | | |
| | L | Posterior-middle<br>Percept: auditory lead > synch | | |
| | R | Middle<br>Percept auditory lead > synch | | |
| | R | Anterior-Middle<br>Percept: Auditory lead vs synch | | |
| **Superior temporal gyrus** | L | Detect order of auditory syllables vs gender detection.<br>Syllable order > Gender ID | Auditory speech syllables | Moser et al., 2009 |
| **Middle temporal gyrus** | L | Detect order of auditory syllables vs gender detection.<br>Syllable order > Gender ID | Auditory speech syllables | Moser et al., 2009 |
| **Middle Frontal Gyrus** | L | Detect order of auditory syllables vs gender detection.<br>Syllable order > Gender ID. | Auditory speech syllables | Moser et al., 2009 |
| **Inferior frontal gyrus** | Bi | Detect order of auditory syllables vs phoneme ID<br>Syllable order > phoneme ID. | Auditory speech syllables | Moser et al., 2009 |
| | R | TOJ activation > shape discrimination | Visual shapes | Davis et al., 2009 |
| **Temporal parietal junction** | L | TOJ activation > shape discrimination | Visual shapes | Davis et al., 2009 |
| **Inferior frontal lobe** | R | TOJ activation > shape discrimination | Visual shapes | Davis et al., 2009 |
| **Inferior parietal lobule** | L | Detect order of auditory syllables vs phoneme ID Syllable order > phoneme ID. | Auditory speech syllables | Moser et al., 2009 |

*Table 3-2: Areas previously identified in TOJ research*

| Anatomical structure | L/R | Task/Contrasts | Stimuli | Reference |
|---|---|---|---|---|
| **Prefrontal cortex** | R | Physical synch > asynch | AV Speech | Noesselt et al., 2012 |
| | L | Physical Auditory lead > synch | | |
| | Bi | Physical visual lead v synch | | |
| | Bi | Percept: visual lead > synch | | |
| | Bi | Percept: Auditory lead > synch | | |
| **Precuneus** | Bi | Detect temporal order vs gender detection<br>Syllable order > Gender ID | Auditory speech syllables | Moser et al., 2009 |
| **Supramarginal gyrus** | Bi | Detect order of auditory syllables vs gender detection.<br>Syllable order > Gender ID | Auditory speech syllables | Moser et al., 2009 |
| | L | TOJ activation > shape discrimination | Visual shapes | Davis et al., 2009 |
| **Intraparietal sulcus** | Bi | TOJ activation > shape discrimination | Visual shapes | Davis et al., 2009 |
| **Insula** | R | Anterior: physical visual lead > synch | AV Speech | Noesselt et al., 2012 |
| | Bi | Anterior<br>Percept: visual lead > synch | | |
| | Bi | Anterior<br>Percept: auditory lead vs synch | | |
| | | | | |
| **Insula/IFG** | L | Anterior/IFG<br>physical visual lead > synch | AV Speech | Noesselt et al., 2012 |
| | | Anterior/IFG<br>physical Auditory lead > synch | | |
| | | | | |
| **Precentral Gyrus** | R | Physical auditory lead > synch | AV Speech | Noesselt et al., 2012 |
| | Bi | Percept: auditory lead > synch | | |
| **Superior frontal gyrus** | Bi | TOJ activation > shape discrimination | Visual shapes | Davis et al., 2009 |
| **Putamen / Lentiform Nucleus** | R | Detect order of auditory syllables vs gender detection.<br>Syllable order > Gender ID | Auditory speech syllables | Moser et al., 2009 |
| **Thalamus** | n/a | TOJ activation > shape discrimination | Visual shapes | Davis et al., 2009 |

*Table 3-3: Areas previously identified in TOJ research (continued)*

Additional structures such as the cerebellum, superior colliculus and putamen have been reported to be implicated in AV synchrony processing, but have not been reported in AV temporal order research. Processing AV temporal order seems to recruit more additional areas compared to processing synchrony, including the inferior frontal lobe (exclusive of the IFG), supramarginal gyrus, intraparietal sulcus, MTG, Precuneous, Precentral gyrus and thalamus. Overall this is indicative of a set of core structures implicated in temporal processing, which might be complemented by the recruitment of additional anatomical structures forming distinct networks, which differ based on whether the task requires processing of AV temporal order or synchrony.

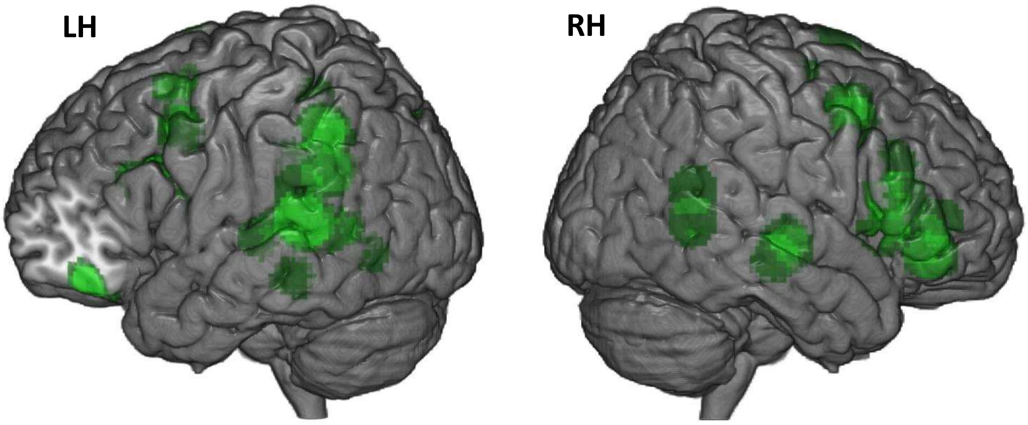


*Figure 3-2: Areas previously identified by fMRI research to be involved in AV and unisensory temporal order processing.*

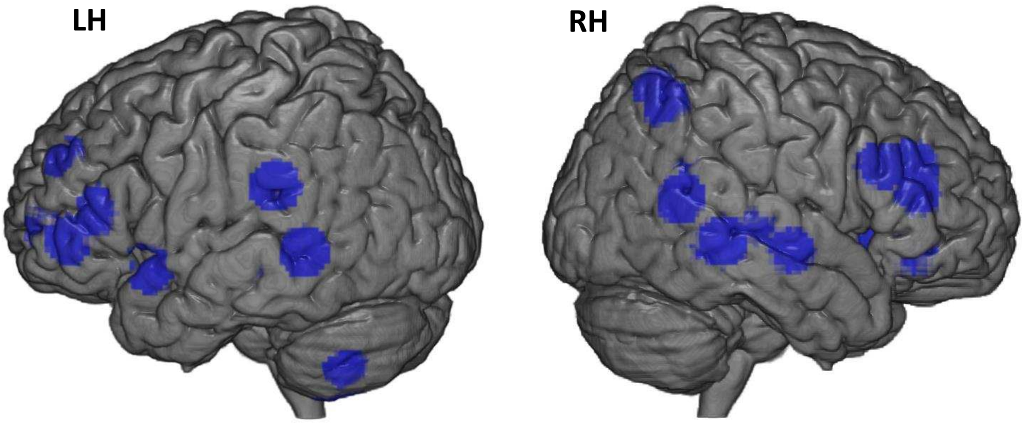


*Figure 3-3: Areas previously identified by fMRI research to be involved in AV synchrony processing.*

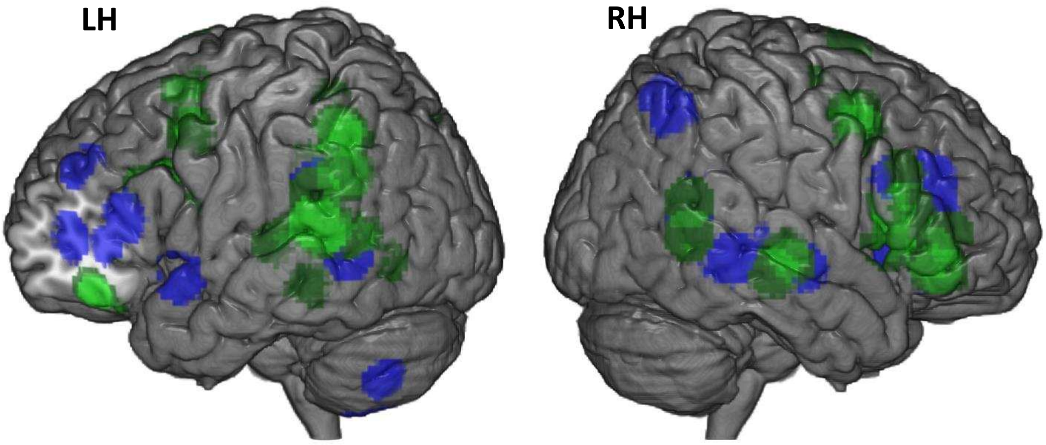


*Figure 3-4: AV Temporal order processing (green) and synchrony processing (blue) areas previously identified in the literature, superimposed for comparison.*

Although the research discussed so far informs on the neural networks that might be generally implicated in AV temporal processing, whether these networks support implicit or explicit AV synchronisation cannot be confidently inferred. One reason for this is the difficulty of separating implicit and explicit temporal processing during online measurement of brain activity, without elaborate control conditions. For example, it is possible that when processing AV synchrony during explicit temporal judgements, an observer is at the same time integrating AV information and consequently also implicitly synchronising. Thus, any comparisons between activity observed during trials in which a temporal judgement is made and control or baseline conditions in which AV integration would not occur, might reveal areas not only related to explicit synchrony processing, but also related to AV integration or implicit AV synchronisation. Because implicit processes are not concurrently measured during these conditions, such activations cannot be partialled out in the analysis. A similar problem is apparent in research which compares activation during the perception of synchrony to activity measured during the perception of asynchrony. At least some increases in activity in synchronous compared to asynchronous conditions could be related to processes underlying the integration of AV information rather than its synchronisation. This is because synchronous AV stimuli are more often than not integrated, and thus some areas reported to be involved in synchronisation might actually be responsible for AV integration.

In a similar vein, it can often be ambiguous as to which areas identified in an investigation play a crucial role in supporting the process being investigated and which areas only support peripheral processes, indirectly associated with the performance of a particular task. One example of this is comparisons between activity observed during trials in which the participant is judging synchrony and during trials during which they are judging some other aspect of the stimulus. Such comparisons could result in activation differences which in fact reflect differences in task difficulty, attentional demands or the degree of certainty with which the observer is making a judgement, rather than differences which reflect synchrony processing. Without elaborate controls conditions, it is very difficult to partial out such activations and often studies are unsuccessful at doing so.

### 3.1.3 NEURAL CORRELATES OF AUDIOVISUAL INTEGRATION IN THE MCGURK EFFECT

The degree to which an observer is on average susceptible to the McGurk effect has been reported to correlate with individual differences in brain activity, specifically with variability in the strength of BOLD responses in posterior STS (Beauchamp et al., 2010). A secondary aim of the current chapter was to investigate whether individual differences in susceptibility to the McGurk illusion might also be related to variability in brain structure. Functional imaging research indicates that the McGurk illusion functionally engages a wide network of cortical structures including temporal, parietal and frontal areas. As can be seen from Table 3-4, Table 3-5 and Table 3-6 on the next page, the exact brain areas which are reported to form this network tend to vary from one study to another. Imaging studies on the neural correlates of the McGurk illusion typically compare BOLD activation for incongruent McGurk AV syllables to activity observed for congruent, incongruent non-McGurk syllables or for both. Other approaches include measuring brain activity to AV stimuli that are classed as likely to be integrated and comparing it to activation that occurs during presentation of AV stimuli that are unlikely to be integrated, for example synchronous versus asynchronous AV stimuli. These studies and their findings will be critically reviewed in the following section.

| Anatomical structure | L/R | Method | Contrast / Analyses | Task | Stimuli | Ref |
|---|---|---|---|---|---|---|
| **Superior Temporal Sulcus** | L | fMRI | AV> nAV masked by nAV-still face(C) | phoneme ID during scan | ba da ga | (Sekiyama, 2003) |
|  | L | PET | AV<nAV masked by nAV-still face(C) | phoneme ID during scan | Ba da ga | (Sekiyama, 2003) |
|  | R | PET | AV<nAV masked by nAV-still face(C) | phoneme ID during scan | Ba da ga | (Sekiyama, 2003) |
|  | L | TMS | TMS vs no TMS | phoneme ID during TMS | Pa ka<br>Ba ga | (Beauchamp et al., 2010) |
|  | L | fMRI | Perceivers vs (>) non perceivers.<br>Positive corr between response & %McG.<br>incongruent > congruent syllables ROI on STS | passive observation in scanner | ba ga da | (Nath & Beauchamp, 2012) |
|  | Bi | fMRI | nonMcG, includes sulcus gyrus, Release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
|  | Bi | fMRI | McG percept > non mcg percept in susceptible group<br>activity in susceptible group > in non susceptible | Detect 'dada' (illusory) | Ba da ga | (Szycik et al., 2012) |
| **Superior Temporal Gyrus** | Bi | fMRI | McG percept > non-McG percept in susceptible group;<br>Activity in susceptible group > in non susceptible | Detect 'dada' (illusory) | Ba ga da | (Szycik et al., 2012) |
|  | Bi | fMRI | release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
| **Middle Frontal Gyrus** | Bi |  | (in susceptible individuals)<br>/pa/ + [ka] > AV Pa<br>/pa/ + [ka] > AV Ka | Passive observation | Apa | (Skipper et al., 2007) |

*Table 3-4: Areas identified in fMRI/PET studies: AV integration using McGurk illusion. // - Auditory, [] - visual*

| Anatomical structure | L/R | Method | Contrast / Analyses | Task | Stimuli | Ref |
|---|---|---|---|---|---|---|
| **Insula** | R | | (in susceptible individuals)<br>/pa/ + [ka] > AV Pa | Passive observation | pa ka | (Skipper et al., 2007) |
| | Bi | fMRI | includes middle temporal gyrus<br>release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
| **Medial Frontal Gyrus** | L | | (in susceptible individuals)<br>/pa/ + [ka] > AV Ka | Passive observation | pa ka | (Skipper et al., 2007) |
| **Superior frontal gyrus** | L | fMRI | (in susceptible individuals)<br>/pa/ + [ka] > AV Pa<br>/pa/ + [ka]> AV Ta | Passive observation | pa ka ta | (Skipper et al., 2007) |
| **Inferior parietal lobule** | Bi | fMRI | (in susceptible individuals)<br>/pa/ + [ka] > AV Pa<br>/pa/ + [ka]> AV Ta | Passive observation | pa ka | (Skipper et al., 2007) |
| | L | fMRI | Incong > cong at synch | ‘ba’ or other? | ava aba<br>at 0 ±400ms | (Jones & Callan, 2003) |
| **Precentral gyrus** | Bi | fMRI | (in susceptible individuals)<br>/pa/ + [ka] > AV Pa | Passive observation | pa ka | (Skipper et al., 2007) |
| | R | fMRI | /ava/ + [aba] > AV aba | ‘Ba’ or other? | ava aba | (Jones & Callan, 2003) |
| **Supramarginal Gyrus** | R | fMRI | (in susceptible individuals)<br>/pa/ + [ka] > AV Pa<br>/pa/ + [ka] > AV Ka | Passive observation | pa ka | (Skipper et al., 2007) |
| | L | fMRI | Non-McG to McG<br>release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
| | R | fMRI | Incong > cong McG at synch | ‘Ba’ or other? | ava aba<br>at 0 ±400ms | (Jones & Callan, 2003) |
| **Anterior superior temporal cortex** | R | fMRI | (in susceptible individuals)<br>/pa/ + /[ka] > AV Pa | Passive observation | pa ka | (Skipper et al., 2007) |

*Table 3-5: Areas identified in fMRI/PET studies: AV integration using McGurk illusion (continued). // - Auditory, [] - visual*

| Anatomical structure | L/R | Method | Contrast / Analyses | Task | Stimuli | Ref |
|---|---|---|---|---|---|---|
| **Lingual Gyrus** | R | fMRI | (in susceptible individuals)<br>/pa/ + [ka] > AV Ka | Passive observation | pa ka | (Skipper et al., 2007) |
|  | R | fMRI | nonMcG to McG<br>release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
| **Cerebellum** | R | PET | AV<nAV masked by nAV-still face(C) | phoneme ID during scan | ba da ga | (Sekiyama, 2003) |
| **Thalamus** | R | PET | AV<nAV masked by nAV-still face(C) | phoneme ID during scan | ba da ga | (Sekiyama, 2003) |
| **Calcarine Sulcus** | Bi | fMRI | nonMcG to McG<br>release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
| **Superior parietal gyrus** | R | fMRI | nonMcG to McG<br>release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
| **Posterior calcarine sulcus** | R | fMRi | nonMcG to McG<br>release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
| **Precentral sulcus** | R | fMRI | nonMcG toMcG<br>release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
| **Occipital pole** | R | fMRI | nonMcG to McG<br>release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
| **Frontomarginal sulcus** | L | fMRI | release from adaptation | congruency task in scanner | pa ka | (Benoit et al., 2010) |
| **Inferior frontal Gyrus** | L |  | (in susceptible individuals)<br>/pa/ + [ka] > AV Pa<br>/pa/ + [ka]> AV Ta | Passive observation | pa ka ta | (Skipper et al., 2007) |

*Table 3-6: Areas identified in fMRI/PET studies: AV integration using McGurk illusion (continued). // - Auditory, [] - visual*

The area most often associated with AV speech integration is the superior temporal sulcus (STS) (Beauchamp, Lee, Argall, & Martin, 2004; Beauchamp et al., 2010; Benoit et al., 2010; Nath & Beauchamp, 2012; Sekiyama, 2003; Wright et al., 2003). The STS has been reported to on average show higher levels of BOLD activation for incongruent McGurk stimuli compared to congruent AV stimuli or incongruent combinations of AV syllables which do not typically elicit the McGurk illusion. Furthermore, individual differences in the strength of STS BOLD responses have been reported to correlate with individual susceptibility to the McGurk effect, measured offline (Nath & Beauchamp, 2012). Nath and Beauchamp presented participants with physically synchronous congruent and incongruent combinations of AV syllables. The incongruent syllables were further divided into two types: combinations designed to elicit the McGurk effect and combinations which were designed to be perceived veridically. Participants were not required to identify which phoneme they believed was uttered on every trial and instead responded to catch trials containing the syllable ‘ma’, to ensure that they attended to the stimuli throughout the scanning session. The analysis was restricted to regions of interest comprising voxels in the STS, Heschl’s gyrus, visual cortex, inferior frontal gyrus and Inferior Precentral Sulcus. The STS was reported to be the only region showing increased activity to incongruent compared to congruent AV stimuli, including both combinations which elicit the McGurk effect and combinations which do not. This area was also found on average to show significantly higher levels of activity in individuals susceptible to the McGurk effect compared to non-susceptible individuals. Participants’ susceptibility was also measured outside the scanner and individual differences in this measure were correlated with individual differences in the signal strength observed in the STS. The signal strength of the STS in each individual was positively correlated to the greater susceptibility to the McGurk effect.

Other temporal areas implicated in the McGurk effect (illustrated in Figure 3.5 on the next page along with all other identified areas) include the superior temporal gyrus (Benoit et al., 2010; Szycik et al., 2012) and the right anterior superior temporal cortex (Skipper et al., 2007). Parietal areas include the Inferior Parietal

Lobule (IPL) (Jones & Callan, 2003; Skipper et al., 2007), Superior Parietal Gyrus (Benoit et al., 2010) and frontal areas comprise the inferior, middle, medial and superior frontal gyri (Skipper et al., 2007) and the frontomarginal gyrus (Benoit et al., 2010).

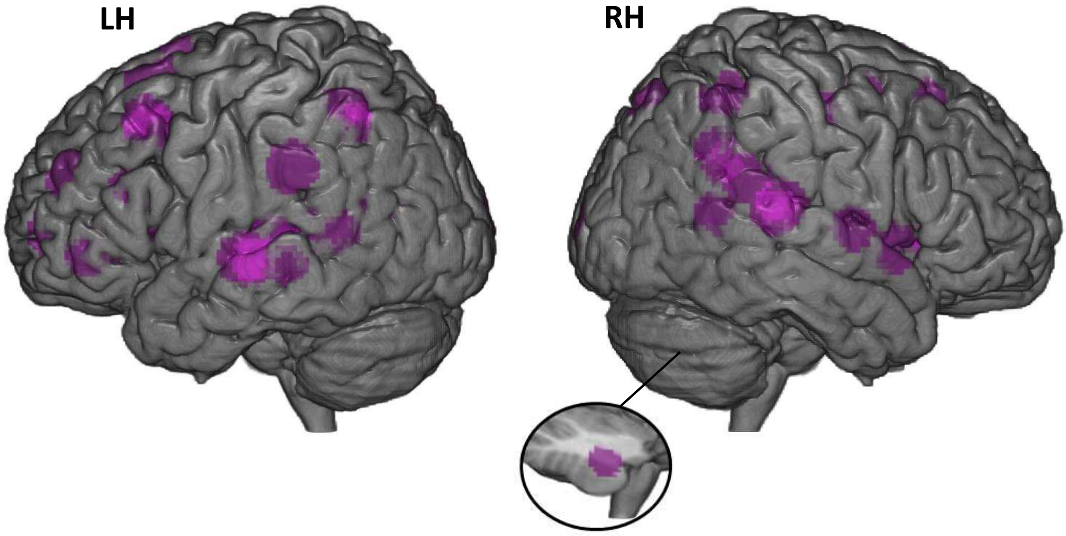


*Figure 3-5: Areas previously identified in fMRI studies to show greater activity during AV McGurk Integration*

Benoit et al. (2010) used a release-from-adaptation paradigm to investigate the neural substrates of the McGurk illusion, in which they measured BOLD responses which correlated with a release from adaptation to congruent syllables. The release from adaption was induced by presenting multiples of congruent AV stimuli, followed by a final pair of incongruent AV syllables. Two types of incongruent AV syllables were used, one combination which was designed to elicit the McGurk effect and another designed to be perceived veridically as incongruent. Whilst being scanned, participants performed an audiovisual congruency task on each of the final pair of stimuli. When participants identified McGurk inducing AV stimuli as congruent they were presumed to have experienced the McGurk illusion. It is unclear as to whether the within group comparisons between McGurk and non-McGurk trials only included McGurk inducing trials in which participants reported that the stimuli were congruent, or whether all trials from this condition were used. The average change in BOLD activity from congruent to incongruent syllables was compared between the two incongruent conditions as well as against activity measured during a baseline condition in which participants viewed a still face. A network comprising the left STS, right insula, right IPS and bilateral primary visual cortex showed more activity in the incongruent McGurk trials compared to incongruent non-McGurk trials. One comparison which would have been useful here, but which was not carried out is a contrast between McGurk inducing trials in which participants reported the syllables as congruent and in which they reported them as incongruent. This contrast might have revealed areas whose contribution give rise to the McGurk percept, with the added benefit of keeping all stimuli properties constant across the comparison. The authors did however run correlations between susceptibility to the McGurk effect and signal change in the areas identified by the group analysis to be involved in the McGurk illusion. They found that the BOLD signal change observed during release-from-adaptation was negatively correlated to the behavioural measures of McGurk susceptibility collected throughout the scanning session. In other words, when the information conveyed by the congruent AV syllables was the same as that conveyed by the illusory percept, the more susceptible a participant was on average to the McGurk illusion, shown by an overall lower likelihood of noticing incongruence, the less likely they were to show a BOLD signal increase indicative of a release from adaptation. These correlations were observed in the left STS, left Heschl's Gyrus, left STG, right Precentral gyrus and the left insula.

Using a more direct way of measuring the McGurk illusion than that used by Benoit et al. (2010), Szycik et al. (2012) presented participants with congruent and incongruent AV syllables, and asked them to perform a phoneme identification task during scanning. In the analyses, the authors also only used McGurk trials during which participants experienced the McGurk effect. Splitting participants into two groups, as a function of whether they were susceptible to the McGurk illusion or not, allowed for comparisons of brain activity between these two groups as well as between the different congruency conditions within the susceptible group. Szycik et al. performed several contrasts, including within and between the two groups. Irrespective of the observers' perceptual experience, significantly higher levels of bilateral STG activation was observed for AV syllables designed to elicit the McGurk effect, in susceptible individuals compared to the non-susceptible group. Within the susceptible group only, significantly greater activation was observed within and around the STS and in the insula for incongruent McGurk syllables compared to congruent ones. Furthermore, contrasts between activity measured during trials in which incongruent AV stimuli elicited the McGurk effect and during trials in which they did not revealed increased activity in clusters within the STG bilaterally. This study therefore indicates that the STG may support the process which underlies the perceptual experience elicited by the McGurk illusion, because this area showed greater activity as a function of susceptibility group and as a function of whether the illusion was experienced by susceptible individuals or not. The observation that regions of the STS and insula showed increased activity as a function of congruency but not as a function of perceptual experience suggest that these structures may not be directly implicated in the perceptual outcome of the McGurk illusion, but that instead they may be involved in more peripheral processes such as detecting incongruence in AV information.

Instead of performing comparisons between susceptible and non-susceptible individuals, Skipper et al. (2007) excluded non-susceptible individuals from their fMRI analysis. They presented participants with AV congruent and with incongruent syllables designed to elicit the McGurk illusion, and with the unisensory components of these stimuli. Participants were not required to report their perceptual experience and were only passively viewing the stimuli inside the scanner. A phoneme identification task performed outside the scanner was used to exclude non-susceptible participants from the fMRI analysis. Activity resulting from exposure to congruent AV syllables was compared against activity resulting from exposure to incongruent McGurk syllables. Incongruent syllables could either share the visual component with the congruent AV syllable, so for example incongruent combination [ka] and /pa/ were compared to the congruent AV ‘ka’, or they could share the auditory component, where for example [ka] and /pa/ were compared to the AV syllable ‘pa’. Contrasts were also run between McGurk syllables and congruent AV syllables which shared the auditory percept, given that the McGurk effect was experienced. Areas showing greater activation to incongruent McGurk syllables relative to congruent AV syllables consisted of the left IPL, right anterior superior temporal cortex (STCa), right Supramarginal gyrus, left medial, middle and superior frontal gyri, right insula and right lingual gyrus. Contrasts between activity to McGurk syllables and congruent AV syllables which matched the illusory percept consisted of the left precuneus, post central gyrus, left inferior, middle, medial and superior frontal gyri. In contrast to other McGurk fMRI studies reviewed here, Skipper et al. (2007) did not find that the STS or STG showed greater activity to incongruent McGurk syllables compared to congruent AV syllables, or non-McGurk incongruent syllables. Instead, their analysis highlighted mainly frontal areas to be involved in McGurk integration. Note however that subjective perception was not measured during scanning, so it is unclear whether the illusion was experienced across all McGurk trials.

Jones and Callan (2003) varied the degree to which typical McGurk incongruent AV stimuli would be integrated using AV asynchrony in order to keep the AV stimuli consistent across comparisons between activity for integrated and non-integrated stimuli. Inside the scanner participants were presented with congruent and incongruent combinations of AV speech stimuli, the latter which comprised pairs of phonemes which typically elicit illusory auditory perception. Both types of stimuli were presented at three SOAs: synchronously, with the visual information leading by 400ms and with the visual lagging by the same amount. The participant’s task was to report whether they heard the phoneme that was actually presented, in this case /b/ or another phoneme. A different answer to /b/ would indicate an illusory response. The authors reported that veridical auditory percepts were collapsed across all the asynchrony conditions and correlated with brain activity; however it is unclear as to whether these responses included those from the congruent condition too. This correlation only revealed a positive relationship between veridical responses and increases in activity within the left occipito-temporal junction, which is puzzling as this is a visual motion processing area. According to the authors, the puzzling result could be accounted for by the observation that the majority of veridical responses were observed in the auditory-leading condition. This in turn made it likely that the auditory information which preceded the visual may have modulated activity in the visual cortex. This explanation does not however concur with the way in which the McGurk illusion manifests. In this illusion it is the visual information that modulates auditory perception and not the other way around. Visual information would therefore be expected to modulate activity in the auditory cortex as opposed to auditory information to modulate the visual cortex. Furthermore, it is unclear why correlations were not run separately for each synchrony condition in order to test this theory. Following the same logic, this relationship would be strongest in the auditory-leading condition and possibly present in the synchronous condition due to the faster processing speed of sound relative to light. This relationship would not be expected in the visual leading condition, where the opposite pattern of results might have been observed; that is, modulation of activity in the auditory cortex would have been expected when the visual information came first. In addition, the authors also compared synchronous congruent and incongruent conditions, which revealed that activity in the right Supramarginal gyrus and left IPL was higher for incongruent stimuli compared to congruent. It is unclear as to whether only incongruent trials in which the illusion was experienced were included in this analysis. This in turn makes it difficult to conclude whether this activation is related to experiencing the McGurk, simply processing and/or detecting incongruent AV stimuli, or a combination of all the above.

A limitation of some of the fMRI literature reviewed so far is that in some studies, participants are only required to passively observe stimuli inside the scanner (e.g. Skipper et al., 2007) and their perceptual experience is not measured concurrently with brain activity. As a result, essentially what these studies measure is brain activity correlated with exposure to a particular type of stimulus and not necessarily activation that can be directly linked to the mechanisms which give rise to the McGurk illusion. This is especially a possibility when control conditions are not carefully selected. Consequently, the link between the underlying mechanisms which give rise to the McGurk illusion and the cortical areas which are said to support them is made indirectly because perception is not actually measured. These studies rely upon the assumption that the individual in the scanner will perceive what he or she is supposed to perceive, which is problematic because exposure to typical McGurk stimuli does not necessarily lead to illusory perception on every occasion. As many of the studies included in this review have reported (e.g. Benoit et al., 2010; Nath & Beauchamp, 2012; Skipper, Nusbaum, & Small, 2005; Szycik et al., 2012) and our behavioural data show, individuals are likely to experience the McGurk effect anywhere between zero and 100% of the time. Therefore, an individual will not experience a given perceptual state every time they are presented with the same pair of AV stimuli, despite all variables being kept constant. Areas showing greater activity during the presentation of these stimuli might therefore not necessarily reflect the neural correlates of mechanisms which give rise to the McGurk illusion. The most one can infer is that these areas are responsive to incongruent AV stimuli, and that the networks identified might or might not contain structures related to illusory McGurk perception. Szycik et al.'s (2012) findings for example demonstrate that areas showing greater activity for incongruent AV information will not necessarily also show increased activation for integrated information and therefore emphasize the need for analysis of brain activity, or structure in this case, as a function of perceptual experience. In this study the STS was more active for incongruent AV stimuli compared to congruent, but did not show the same differential activity as a function of whether the illusion was experienced or not; the latter was observed in the STG instead.

In summary, it is difficult to judge which areas are likely to play a central role in the McGurk illusion and which are likely to support other qualitatively different AV processes from fMRI investigations which do not concurrently measure perception and take it into account in their analyses. In order to suggest that a given network of cortical structures underlies the process of AV integration, it is important to ensure that the information presented is actually being integrated by measuring the observer’s perceptual experience. Brain maps of the mechanisms underlying AV integration should then account for this perceptual measure, rather than simply be based on contrasts of activity measured during conditions in which integration is assumed to take place and in which AV integration is assumed not to occur. Studies which have measured brain activity as a function of perceptual experience might seem methodologically sound, but they are prone to a different methodological criticism: they do not take into account individual differences in the optimal AV asynchrony for AV integration.

Chapter 2 demonstrated that the AV asynchrony optimal for AV integration varies widely across individuals. Functional MRI research which has investigated the McGurk effect has however always presented AV streams synchronously. For some individuals, synchrony is indeed optimal for AV integration. However, for other individuals, who in fact formed the majority of our sample, the visual information needs to lead or lag the auditory rather than to be synchronous, in order to ensure that they integrate the information most frequently. It is therefore possible that presenting an individual with AV information which does not meet their optimal temporal criteria for AV integration leads to the recruitment of additional - but not necessarily successful - AV synchronisation mechanisms, or structures which are sensitive to AV streams that are perceived to be asynchronous. This could in turn mean that some of the brain activity observed is related to AV temporal processing and not the integration of the stimuli. Thus, the cortical maps previously reported to be associated with AV integration might actually also comprise structures responsible for AV synchronisation, rather than only those responsible for AV integration. This could explain why the neural correlates identified by AV integration fMRI studies are very similar to those identified by AV temporal processing literature (see Figure 3-6, on the next page). It could also explain why some functional correlations are not always replicated across the literature, and have only so far appeared in single studies.

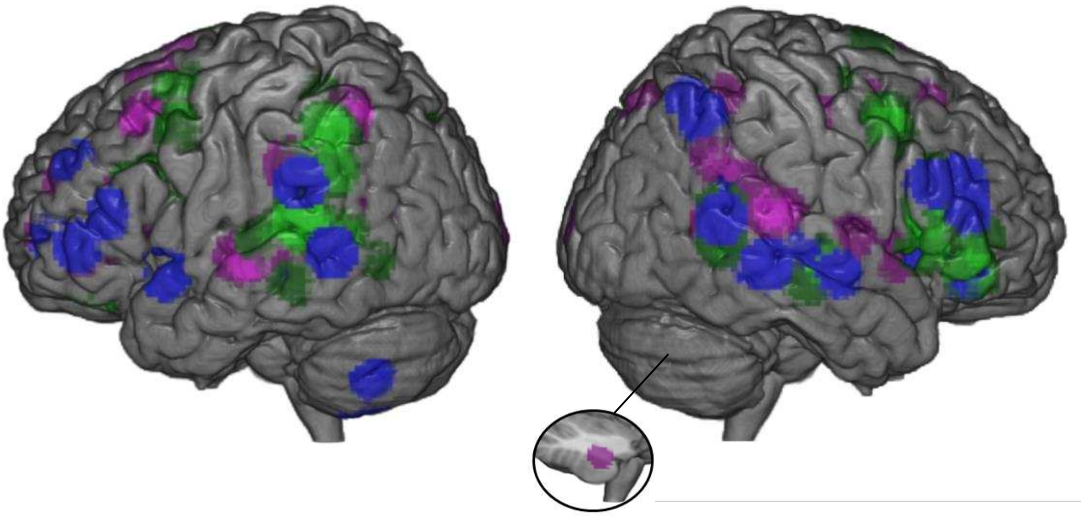


*Figure 3-6: Areas previously identified in previous fMRI literature to show increased activity during McGurk integration (violet), AV temporal order processing (green) and AV synchrony processing (blue)*

The cerebellum, for example is known to be involved in the performance of various temporal tasks such as duration perception and interval timing (Ivry & Spencer, 2004; Mauk & Buonomano, 2004; Meck, 2005). The cerebellum has also been has been reported to show greater activity during both AV synchrony processing (Bushara et al., 2001; Stevenson et al., 2010) and AV integration (Sekiyama, 2003), but only once in the latter. It is possible that this structure might be related to AV synchronisation, or detection of AV asynchrony, but not to AV integration. Greater activity in the cerebellum might have been found on average during AV integration in the study by Sekiyama (2003) either because AV synchrony may not have been the optimal AV relative timing for AV integration for the majority of participants or because the stimuli were on average perceived as asynchronous. As a result, the activity observed in the cerebellum may have originated from additional temporal processes active as a result of exposure to stimuli which are objectively synchronous, but arrive asynchronously in the nervous system. Measuring AV integration as a function of AV synchrony prior to scanning to determine each individual's optimal AV asynchrony for AV integration could be one solution for this, as it would ensure that the stimuli presented will be optimal for AV integration for all participants, and reduce the likelihood that brain activity identified actually reflects additional processes responsive to subjective asynchrony rather than AV integration.

In the current investigation, the measure representing susceptibility to the McGurk illusion is not restricted by individual differences in subjective synchrony. This measure is derived from the entire temporal profile of AV integration. That is, AV integration is measured as a function of AV asynchrony, resulting in a bell shaped curve whose peak represents the highest proportion of McGurk responses observed. This peak can occur at any of the AV asynchronies at which the stimuli were presented. For example, Figure 3-7 on the next page shows two hypothetical profiles of AV integration which illustrates how measuring AV integration only at synchrony may result in measures that do not reflect the observer's susceptibility accurately. Although for clarity, this particular example is hypothetical, it is based on similar observations made on real data (Freeman et al., 2013). At synchrony, participant B seems to integrate AV information more than participant A. However the optimal asynchrony for AV integration for each participant is different. In other words, participant A shows maximum AV integration when the auditory information is presented slightly later than the visual, whereas participant B shows maximum AV integration when auditory and visual information are presented synchronously. The peak of the overall temporal function of AV integration is in fact higher for participant A than it is for participant B but this would not have been apparent if AV integration was measured at synchrony only. Thus, without other factors such as individual differences in the AV asynchrony optimal for AV integration, participant A is actually on average more susceptible to the McGurk illusion than participant B, but this would not have been clear if the entire profile of AV integration had not been measured.

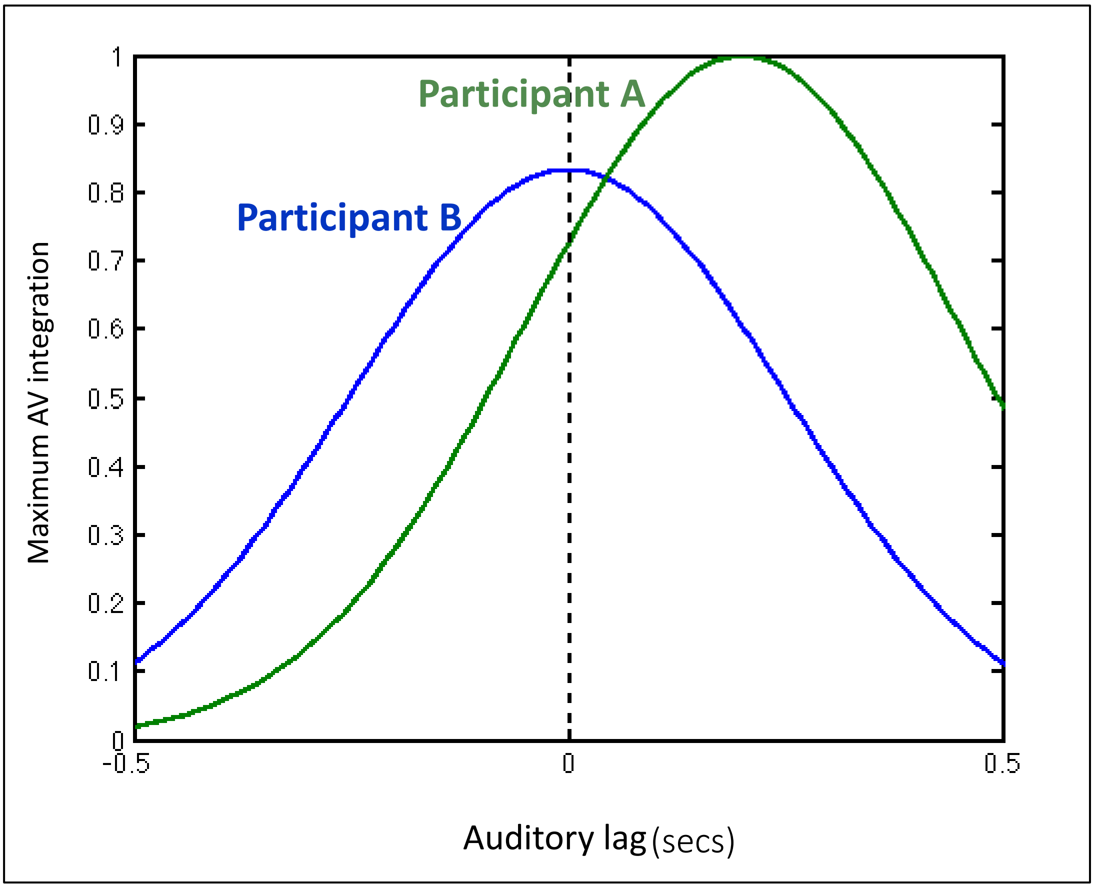


*Figure 3-7: Hypothetical profiles of AV integration: For participants B, AV synchrony is the optimal timing for maximal AV integration. For participant A, the auditory stream needs to be lag the visual in order to elicit maximal AV integration. At AV synchrony, Participant B shows higher a level of AV integration, whereas overall, Participant A shows the highest level.*

In summary of this section, AV integration has been reported to engage a large scale network comprising structures within the frontal, parietal, insular and temporal cortices, the cerebellum and thalamus, some of have been also reported to be recruited during AV temporal processing. The STS is the area most associated with the McGurk illusion. Other temporal areas include the superior temporal gyrus (Benoit et al., 2010; Szycik et al., 2012) and the right anterior superior temporal cortex (Skipper et al., 2007). Parietal areas reported to be recruited during processing of McGurk stimuli are the Inferior Parietal Lobule (IPL) (Jones & Callan, 2003; Skipper et al., 2007), Superior Parietal Gyrus (Benoit et al., 2010) and frontal areas comprise the inferior, middle, medial and superior frontal gyri (Skipper et al., 2007) and the frontomarginal gyrus (Benoit et al., 2010).

The exact roles of the components of these networks is however unclear, due to the possibility that they may contain structures related to the perceived relative timing of AV information. This might be because whilst the optimal AV asynchrony for AV integration varies widely across individuals and for most individuals AV synchrony is not the optimal AV relative timing for AV integration, AV stimuli in the studies reviewed are always presented at synchrony. Performing one task may also activate a network of multiple areas, each serving other, complementary processes. For example, the perceptual outcome of AV integration is contingent upon the temporal correspondence between the auditory and visual components. We know this because AV synchrony affects the likelihood that a pair of auditory and visual stimuli will be integrated. In order to integrate AV information, its temporal characteristics must therefore be evaluated, which may mean that to two separate processes are active during AV integration: AV synchronisation and AV integration. Although these two processes are likely to be equally important to the overall process of AV integration, when mapping the neural correlates of AV integration it is still important to distinguish areas which support AV integration from those supporting complementary processes. Traditional fMRI approaches often rely on correlations between brain activity and online performance, making it difficult to make this distinction because multiple processes necessary for a given task may be active concurrently. Even if these processes occur at different times, for example, if synchronisation occurs before integration, fMRI lacks the temporal discriminability necessary to tease them apart. Finally, the observer’s perceptual experience is not always taken into account in AV integration fMRI studies, which leads to ambiguity as to whether all the components of the neural networks identified by such studies are directly involved in AV integration, or whether they support other peripheral AV processes such as the detection of AV incongruence.

### 3.1.4 QUANTIFYING PERFORMANCE IN IMPLICIT AND EXPLICIT TEMPORAL PROCESSING

To maximise the likelihood that structural and behavioural correlations in this chapter reflect potential differences or similarities between the underlying temporal mechanisms and neural substrates of implicit and explicit temporal processing, as opposed to differences in the way in which the data was handled prior to analysis, the method used here to model the data was based on Yarrow et al.'s (2011) Two-Criterion SJ model. This particular method was firstly chosen because it uses cumulative Gaussian functions to fit SJ data which can be fitted to AV integration data in the same way. The TOJ data can then also be fitted with a single cumulative Gaussian. Critically, comparable measures can be derived in the same way from the profiles of AV integration, temporal order judgements and synchrony judgements.

Secondly, the noisy criteria model may be a more appropriate method for fitting SJ data. When judging temporal order, participants use only one criterion for deciding whether the auditory stream preceded or followed the visual stream. The decision of whether the sound occurred first or second will depend on which side of this single criterion the participant’s perception falls on. The slope of the cumulative Gaussian function fitted to TOJ data, which represents temporal sensitivity, does not depend on this criterion. However, when judging whether a stimulus pair is synchronous, two criteria are used, one for visual leading and another for visual lagging stimuli. If the observer’s perception of AV relative timing for a given stimulus pair falls within these two criteria, then a ‘synchronous’ response will be made and if it falls outside the two criteria, then the stimulus will be judged as asynchronous. When estimated in the traditional way as depicted in Figure 3.8, the JND from SJs relies heavily on where the participant has placed these two criteria, thus what is often regarded as a measure of temporal sensitivity could just as likely reflect response bias. Computing SJ JNDs using the traditional method is also problematic for comparing temporal discriminability across SJs and TOJs. Any differences found between the two measures within one individual could reflect that the JND from SJs relies on where the observer sets their decision criteria and that the JND from TOJs does not.

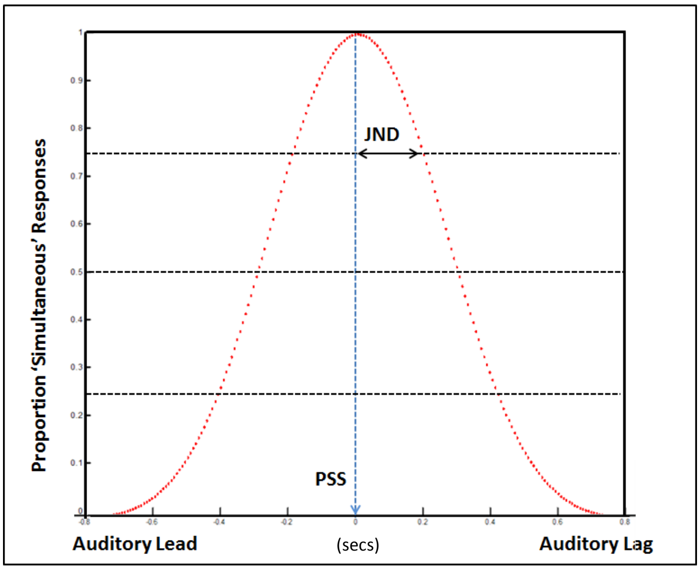


*Figure 3-8: Hypothetical SJ data, plotted as a function as AV asynchrony with the JND estimated by halving the distance between the two SOAs at which the participant responded 'synchronous' 75% of the time.*

For the above reason, fitting two cumulative Gaussians to SJ data and subtracting one from the other to obtain a bell-shaped curve (as seen in Figure 3.9 on the next page) may be more appropriate for fitting SJ data in order to estimate temporal sensitivity. Measuring the slopes of the cumulative Gaussians and computing an average of the two will produce a JND measure which is less dependent on response criteria settings and more analogous to that of TOJs, allowing for direct comparison of temporal discriminability across the two tasks. This is because the standard deviation of the cumulative Gaussians does not rely on the distance between the two criteria as much as the JND measure does.

The model assumes that in order to decide whether a given stimulus pair was synchronous or not participants need to adopt two criteria, an auditory-leading boundary and an auditory-lagging one and that these criteria can themselves be noisy. Within these criteria, participants choose a 'synchronous' response and outside of the criteria they choose an 'asynchronous' response. The second probability function is subtracted from the first, resulting in a normal distribution curve typical of synchrony judgements (see Figure 3.9). This type of fit will also allow a measurement of the window of AV synchrony, which can be derived from the distance between the means of the cumulative Gaussians. What the window of synchrony actually represents can be ambiguous because like the SJ JND derived in the traditional way, it is also likely to be affected by both temporal sensitivity and decision criteria.

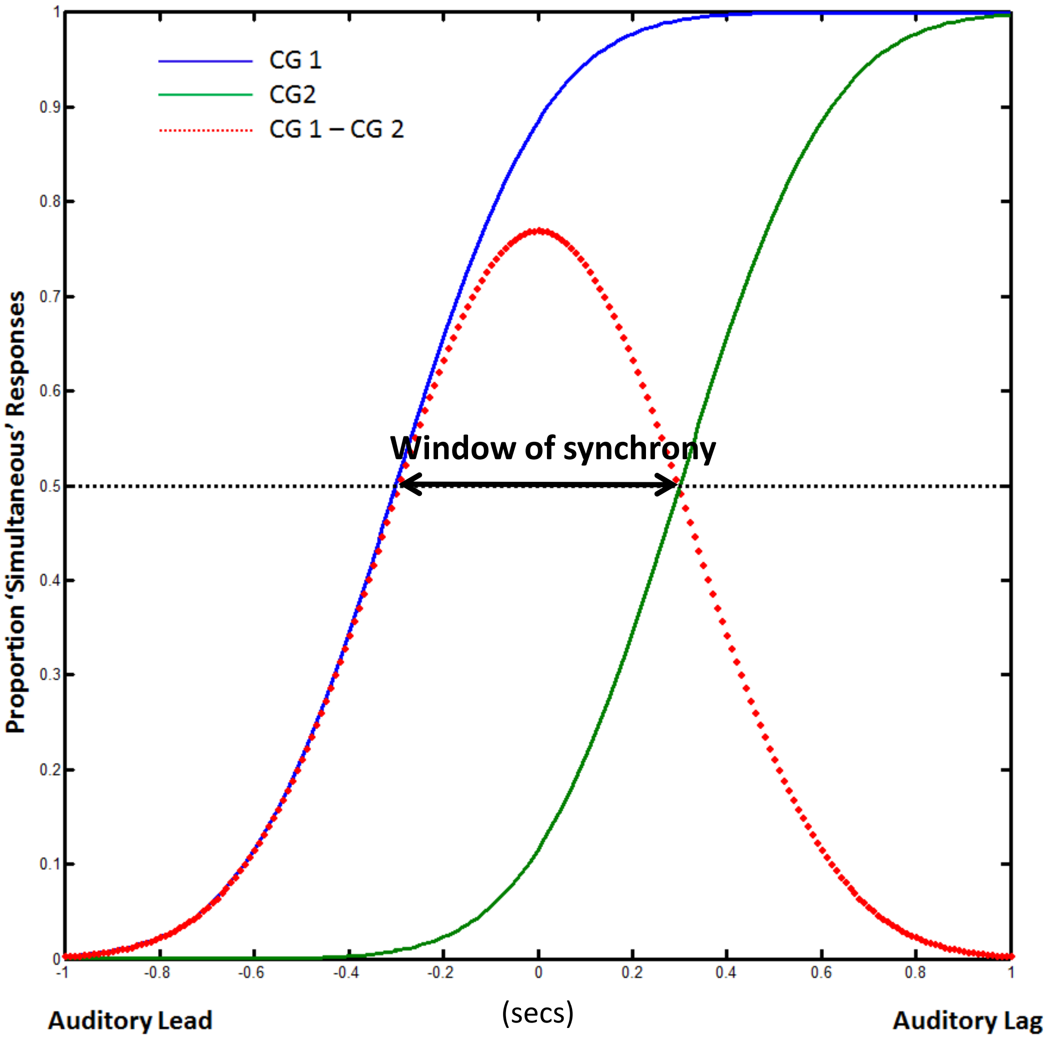


*Figure 3-9 Hypothetical SJ data fitted with Noisy criterion fits (Yarrow et al., 2011) for two-alternative forced-choice synchrony judgements, based on Signal Detection framework: Two Cumulative Gaussians (CG), representing the two criteria adopted by observers when deciding whether a given stimulus pair was synchronous or not. CG 1 (Blue) represents the low criterion and CG 2 (green) represents the high criterion, that is the auditory-leading and auditory-lagging criteria within which participants choose a 'synchronous' response and outside of which they choose an 'asynchronous' response. CG 2 is subtracted from CG1, resulting in a psychometric function whose width, height and slope can be estimated (red). The 'Win' parameter, labelled '**Window of synchrony'** is derived from the distance between the means of the cumulative Gaussians, whereas the **SD** measures are estimated from the average of the standard deviation for each of the cumulative Gaussians.*

The standard deviation of the cumulative Gaussian (or the average standard deviation of the two cumulative Gaussians that are fitted to synchrony and AV integration data) is an estimate of its slope and represents the ability to discriminate between perceived synchronous and asynchronous AV stimuli. This measure is similar to the JND derived from TOJs, but has the added benefit that, when derived from SJ and AV integration data, it does not depend on the distance between the two sides of the distribution curve and only represents the slope. In other words, it represents how quickly participants switch from one response to the other as a function of AV asynchrony. The second measure of performance, which is only derived from the SJ and AV integration data, quantifies the distance between the two cumulative Gaussians, i.e. the width of the temporal profiles of AV synchrony perception and AV integration, and represents the degree to which asynchronous AV stimuli are judged to be synchronous or integrated.

Each measure is therefore represented by a single parameter which is extracted in the same way from each the temporal profiles of TOJ, SJs and of AV integration, and categorised as implicit or explicit temporal performance according to the task it was derived from. The measure of temporal discriminability (SD) does not depend on the PSS as the function is free to shift along the ‘x’ axis and its SD is unaffected. Mathematically, the SD measure is also independent from the other parameter derived from temporal profiles of AV synchrony and AV integration and the window of synchrony (SJ Win) and that of AV integration (McG Win). For these reasons, the measures used here are very specific in terms of what they represent, thus correlations between brain structure and this parameter are less likely to reveal structures related to peripheral processes.

These parameters are based on explicit and implicit timing judgement data which are obtained in a dual-task, meaning that they are derived from judgements made on the same stimuli, at the same time. Any differences in the brain-behaviour relationships observed across implicit and explicit tasks are thus less likely to reflect changes in attentional demands or stimulus properties, and more likely to reflect subtle differences or similarities between these potentially distinct temporal processes. Altogether the benefits of the methodology used here mean that it could reveal, more readily, the areas which correlate with specific aspects of performance in a task, providing a less ambiguous brain-behaviour association.

The SD and Win measures used here are likely to reflect the performance of processes intended to resolve computational complexity that arises as a result of external and internal inconsistency in the relative timing of AV information. There is a possibility that the more resources a given areas has, in this case a larger volume of grey matter, the better equipped it is to perform the neural computations necessary for the mechanisms it supports. For example, one theory is that AV relative timing is estimated from the distribution of activity derived from populations of neurons, each responsive to different AV temporal delays (Roach, Heron, Whitaker, & McGraw, 2011). A distribution of signals produced by larger populations of neurons would be subject to less noise, and thus more reliable, leading to better discrimination of AV synchrony or temporal order. Thus, individuals with larger grey matter volume in areas responsible for temporal processing might be more likely to have higher sensitivity when discriminating AV asynchrony from synchrony, and smaller windows of AV integration or subjective synchrony.

Correlating specific parameters derived from temporal functions to brain structure takes advantage of the possibility that individuals can differ from one another in more than one aspect of AV temporal processing ability, and that performance in different aspects of temporal processing may vary independently. The behavioural measures used throughout the current thesis each reflect different isolated aspects of AV integration and timing. For example, in the current chapter, measures which represent the temporal specificity of AV integration are mathematically independent of the degree to which an individual is susceptible to the McGurk illusion as well as of the optimal AV asynchrony for AV integration. In the same vein, the measures of AV temporal discrimination do not rely on subjective synchrony as the SD of the function is estimated independently from the mean of the function. Thus, any clusters of brain voxels identified to correlate with AV temporal discrimination measures will reflect structures related to discrimination ability, and not to how close to physical synchrony the observer’s subjective perception of AV synchrony is. This is important as the latter may be dependent on the morphology of other areas, for example within unisensory cortices, as well as different structural characteristics of the brain such as the strength of connectivity between different structures.

### 3.1.5 VOXEL-BASED MORPHOMETRY

The majority of research discussed in the current chapter has attempted to identify the neural networks underlying the processing of AV synchrony, temporal order and AV integration which are common across groups of individuals. The current investigation is not concerned with mapping the neural sites of AV temporal and integration mechanisms which are similar across groups of individuals to the brain but rather, it aims to identify structures that might underlie differences in the performance of AV integration and temporal processing mechanisms across individuals. Furthermore the study aims to address the more specific question of whether a distinction can be made between the underlying mechanisms of implicit and explicit temporal processing. As the previous literature review has exposed, using functional imaging methods to answer this question could be possible, but only with elaborate control conditions and contrast analyses, which are absent in the existing literature. This chapter therefore used Voxel-Based Morphometry (VBM) to correlate individual differences in AV temporal processing and AV integration with individual variation in brain structure, to address whether implicit and explicit AV temporal processing might be supported by common or by distinct temporal mechanisms. VBM used in conjunction with individual differences is a relatively novel approach (Kanai et al., 2010; van Gaal et al., 2011). Because it makes use of the relationship between behaviour and brain structure, which is a static characteristic of the brain compared to a dynamic one such as BOLD responses, elaborate control conditions and contrasts that would be necessary in an fMRI investigation, are not required. Using VBM, the chapter examines whether performance in implicit and explicit temporal processing is related to grey matter volume in common or distinct anatomical areas.

In summary, under assumptions of unity, individual differences in AV temporal processing performance should correlate positively across implicit and explicit AV timing tasks as well as with grey matter volume in common clusters. Alternatively, if performance is not consistent across the different AV temporal tasks and correlates with distinct clusters, it might be an indication that qualitatively different temporal processes are supported by distinct underlying structures. Under the assumption that larger populations of neurons would produce distributions of signals subject to less noise, it was expected that better performance should correlate with larger density of grey matter volume. To test whether implicit and explicit AV synchronisation are supported by distinct or common neural mechanisms, and whether these areas are in turn common or distinct from areas correlated with AV integration, local grey matter volume at individual voxels of the brain was measured using VBM and correlated to specific parameters of psychometric functions fitted to explicit and implicit AV temporal judgement data.

# 3.2 EXPERIMENTS 3.1 AND 3.2

## 3.2.1 METHODS

### 3.2.1.1 SUBJECTS

Twenty-seven neurologically healthy young subjects (18-28 years, mean 22) took part in the experiment. Data from four further participants were excluded, due to poor performance, resulting in implausible estimates of subjective timing >300ms asynchrony, outside the typical range for multisensory integration (Vatakis, Ghazanfar, et al., 2008; Vatakis & Spence, 2007) and indicative of poor quality data and unreliable function fits.

### 3.2.1.2 STIMULI & APPARATUS

The same stimuli were used as in Experiment 2.1 of Chapter 2.

### 3.2.1.3 PROCEDURE

The same procedure was used as in Experiment 2.1 of Chapter 2.

### 3.2.1.4 ANALYSIS

#### 3.2.1.4.1 Synchrony Judgements and Phoneme ID

For the SJ task, the proportion of 'synchronous' responses was plotted as a function of asynchrony for each of the 4 conditions. The average proportions across the 4 conditions were also plotted. For the Phoneme ID task, the proportion of illusory responses was plotted as a function of AV asynchrony for each of the two incongruent conditions, along with the average proportions. Each set of data was then fitted with a difference of two asymmetric cumulative Gaussians function (Yarrow et al., 2011), from which the parameters used in the analysis were extracted and averaged out across the conditions and their average curves (for an example of this see Figures 3-10 and 3-11 on the next two pages). These parameters were the average standard deviation (SD) of the cumulative Gaussians and the difference between the means of the Gaussians in terms of AV asynchrony. For the McGurk data, the peak of the resulting function, labelled here 'McG Max' was also extracted from the function and represents susceptibility to the McGurk Illusion. The SD parameter is an estimate of the average slope of the two cumulative Gaussians and represents how abruptly individuals switch from one response to another as a function of AV asynchrony. In other words, the SD is a measurement of the sensitivity with which AV asynchrony is discriminated from AV synchrony, and small measures of the SD represent higher sensitivity. The difference between the means of the cumulative Gaussians represents an absolute value of the width of the window of AV synchrony perception and the temporal window of AV information. In other words, the Win parameter represents the temporal specificity of AV synchrony perception and AV integration, respectively. Small measures of the Win parameter represent small windows of AV integration and synchrony perception. Lastly, high measures of ‘McG max’ represent high susceptibility to the McGurk illusion.

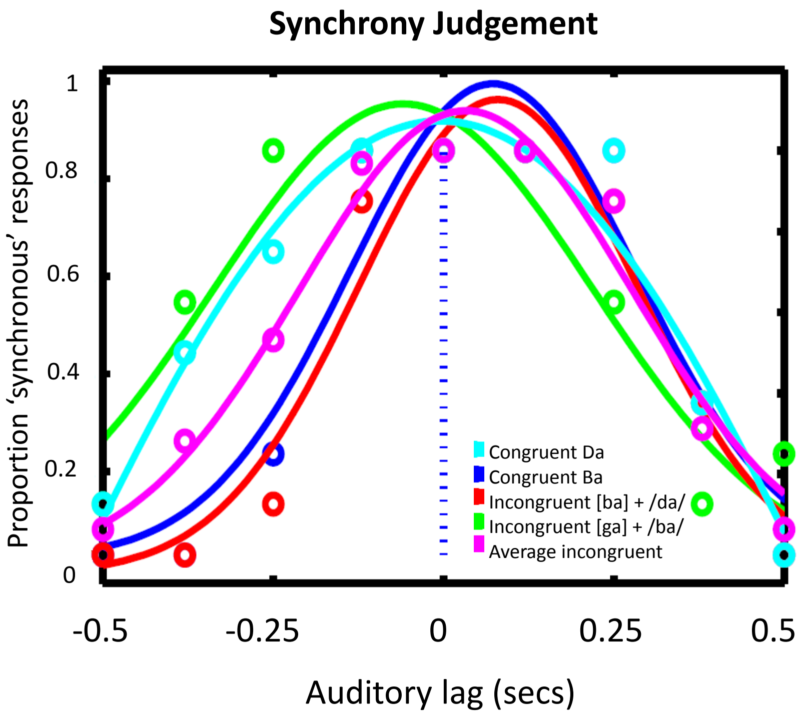


*Figure 3-10: Examples of Noisy Criterion functions fitted to SJ data.*

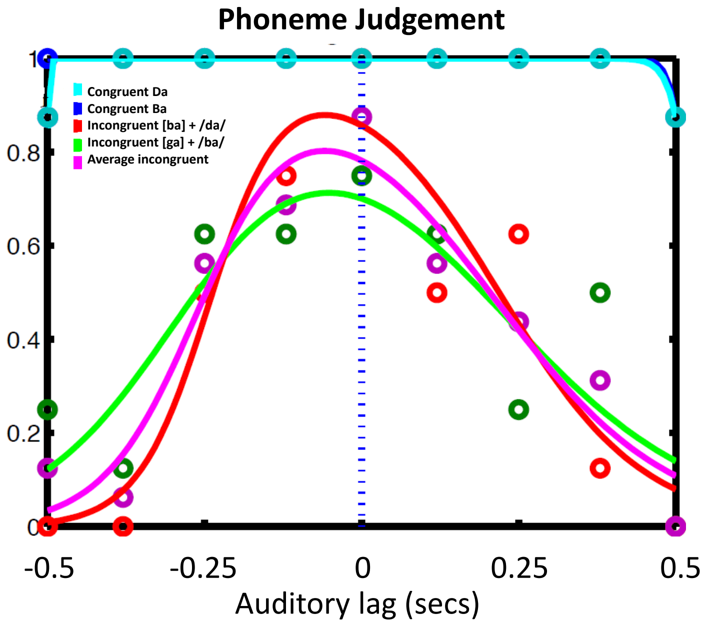


*Figure 3-11: Example of Noisy criterion function fitted to phoneme ID (McGurk) data.*

#### 3.2.1.4.2 Temporal order judgements

The proportion of ‘sound second’ responses was plotted as a function of auditory lag and fitted with a cumulative Gaussian function (see Figure 3.12 on the next page for example). The SD of the function was extracted and represents how abruptly individuals switch from ‘sound first’ responses to ‘sound second’ responses as a function of AV asynchrony. The SD is a measurement of the sensitivity with which AV synchrony is discriminated from AV asynchrony.

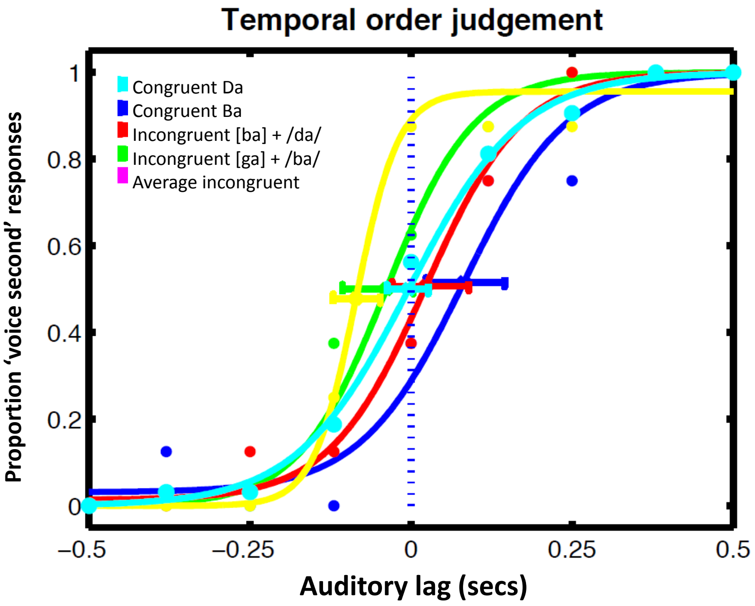


*Figure 3-12: Example of TOJ data fitted with a Cumulative Gaussian function*

### 3.2.1.5 IMAGE ACQUISITION

MR images were acquired on a 1.5-T Siemens Sonata MRI scanner (Siemens Medical). High-resolution anatomical images were acquired using a T1-weighted 3D Modified Driven Equilibrium Fourier Transform sequence (repetition time = 12.24ms; echo time = 3.56ms; field of view = 256 x 256mm; voxel size = 1 x 1 x 1mm).

### 3.2.1.6 VBM PRE-PROCESSING AND ANALYSIS

T1-weighted MR images were first segmented for grey matter and white matter using the segmentation tools in Statistical Parametric Mapping 8 (SPM8, ttp://www.fil.ion.ucl.ac.uk/spm). Diffeomorphic anatomical registration was then performed through exponentiated lie algebra in SPM8 for inter-subject registration of the grey matter images (Fleming, Weil, Nagy, Dolan, & Rees, 2010). To ensure that the total amount of grey matter was conserved after spatial transformation, transformed images were modulated by the Jacobian determinants of the deformation field. The registered images were then smoothed with a Gaussian kernel of 12 mm full-width half-maximum and were then transformed to Montreal

Neurological Institute stereotactic space using affine and nonlinear spatial normalisation implemented in SPM8.

Statistical analyses were performed with SPM8 (Statistical Parametric Mapping, Welcome Department of Cognitive Neurology, University College London, London, UK, http://www.fil.ion.ucl.ac.uk.spm) and the non-stationary (NS) toolbox.

## 3.2.2 RESULTS

### 3.2.2.1 BEHAVIOURAL RESULTS

Correlations were run to explore the relationship between implicit and explicit AV temporal processing. Scatter plots are illustrated in Figure 3.13 and statistics are summarised in Table 3.7, both on the next page. Performance in discriminating AV temporal order (TOJ SD) correlated significantly and positively with the ability to discriminate AV synchrony from asynchrony (SJ SD) [$r_s(27)=.759$, $p<.0005$] within participants (Figure. 3-13 a.), meaning that individuals who are good at discriminating between synchronous and asynchronous AV stimuli are also good at discriminating the temporal order of AV events. The window of AV synchrony (SJ Win) derived from the SJ task correlated positively with the ability to discriminate AV temporal order (TOJ SD) [$r_s(27)=.626$, $p<.0005$] (Figure. 3-13 b.) and moderately with the ability to discriminate AV synchrony in SJs (SJ SD) [$r_s(27)=.440$, $p=.022$] (Figure. 3-13 c.). Individuals who are better at discriminating temporal order and synchrony of AV events tend to have smaller windows of synchrony.

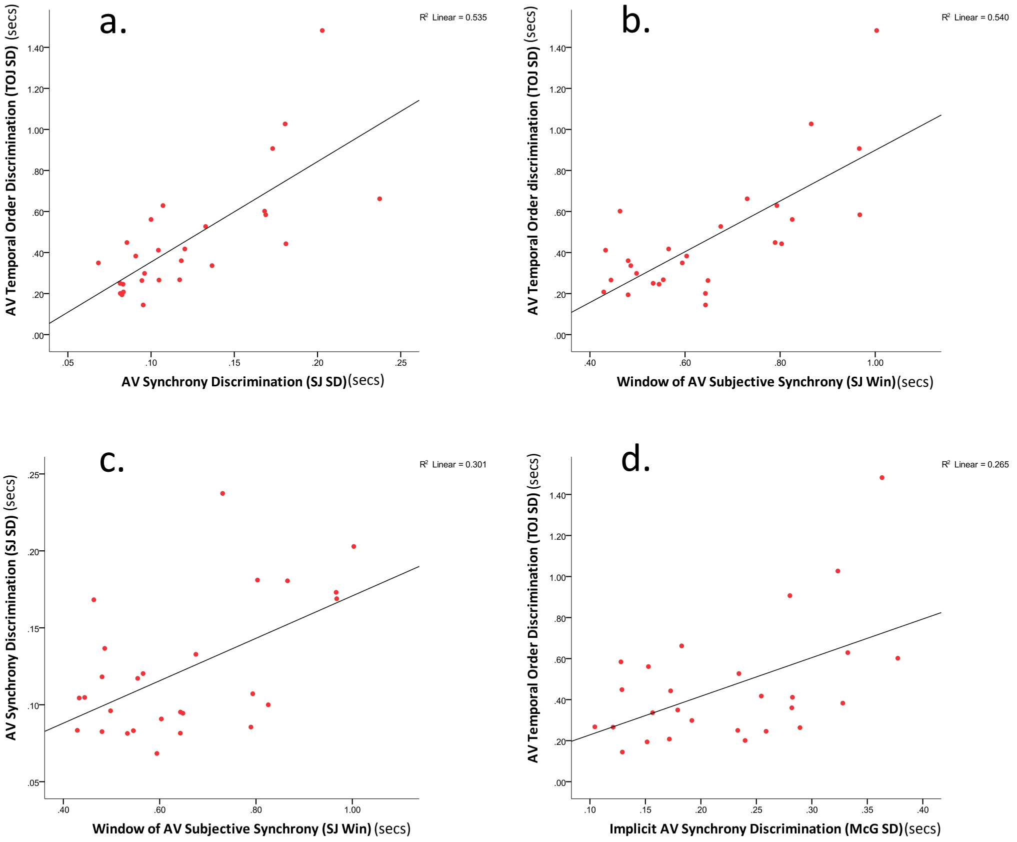


*Figure 3-13: Scatter plots of significant correlations between the temporal processing performance measures.* ***a.*** *significant positive correlations between TOJ SD and SJ SD.* ***b.*** *significant positive correlations between TOJ SD and SJ Win.* ***c.*** *significant positive correlations between SJ Win and SJ SD.* ***d.*** *significant positive correlations between TOJ SD and McG Win**.***

| | | **Explicit** | | | **Implicit** | |
|---|---|---|---|---|---|---|
| | | **TOJ SD** | **SJ SD** | **SJ Win** | **McG SD** | **McG Win** |
| **Explicit** | **TOJ SD** | **---** | | | | |
| | **SJ SD** | $r_s(27)=.759$,<br>$p<.0005$ | | | | |
| | **SJ Win** | $r_s(27)=.626$,<br>$p<.0005$ | $r_s(27)=.440$,<br>$p=.022$ | | | |
| **Implicit** | **McG SD** | $r_s(27)=.401$,<br>$p=.038$ | $r_s(27)=.234$,<br>$p=.241$ | $r_s(27)=.103$,<br>$p=.611$ | | |
| | **McG Win** | $r_s(27)=.046$,<br>$p=.821$ | $r_s(27)=.191$,<br>$p=.340$ | $r_s(27)=-.052$,<br>$p=.797$ | $r(27)=.366$,<br>$p=.061$ | |

*Table 3-7: Summary of significant (green) and non-significant (red) behavioural correlations of Chapter 3.*

A one way repeated measures ANOVA revealed that there was a significant main effect of type of task on the SD parameter [$F(2,52)=33.75$, $p<.0005$]. On average, the TOJ SD (M=46, SD=0.05) was significantly larger than SJ SD (M=0.12, SD=0.01) ($p<.0005$, Bonferroni adjusted) measure, implying that temporal order is more difficult to judge compared to AV synchrony.

Performance in implicit discrimination of AV synchrony (McG SD) correlated moderately with the ability to discriminate AV temporal order [$r_s(27)=.401$, $p=.038$] (Figure. 3.13 d.) but did not significantly correlate with performance in discriminating synchronous from asynchronous AV stimuli (SJ SD) [$r_s(27)=.234$, $p=.241$]. Individuals with high sensitivity to AV asynchrony during AV integration are significantly more likely to have higher sensitivity to the temporal order of AV events during explicit temporal order judgements, but they are not significantly more likely to have high sensitivity to asynchrony during explicit synchrony judgements. Discrimination of AV synchrony was on average significantly poorer when it was performed implicitly (M=0.22, SD=0.08) compared to explicitly (M=0.12, SD=0.05) [$p<.0005$, Bonferroni adjusted].

The temporal window of AV integration (McG Win) did not correlate significantly with any of the explicit timing measures (see table 3-7 on previous page for statistics), and it did not correlate with the ability to implicitly discriminate between synchronous and asynchronous AV stimuli [$r_s(27)=.331$, $p=.091$]. The window of AV integration was also significantly smaller on average than the window of AV synchrony [$t(26)=-5.63$, $p<.0005$].

## 3.2.2.2 VBM RESULTS

### 3.2.2.2.1 ROI Analysis: Superior and Middle temporal cortex

Structural data were first analysed using a region of interest (ROI) analysis. A mask for the superior and middle temporal cortex was constructed using the MarsBar toolbox in SPM. This portion of the temporal cortex was chosen because activity in this area has been reported most frequently to be related to processing of AV relative timing and has been reported to correlate with individual differences in AV integration (Nath & Beauchamp, 2012). The aim of this ROI analysis was therefore to explore whether individual differences in AV integration and timing are correlated with structural variability in this area. Specifically, the first aim of the analysis was to reveal whether individual differences in performance of implicit and explicit AV temporal processing correlates with distinct subregions within the superior and middle temporal cortex. Furthermore, the analysis also aimed to converge on and complement the finding that complementary processes within AV integration are related to distinct neural substrates using structural correlations (Stevenson, vanDerKlok, Pisoni, & James, 2011; Stevenson, Altieri, Kim, Pisoni, & James, 2010)

Separate ROI analyses (see Figure 3.14 on the next page for illustration of the mask used) were performed for each measure using the SPM8 extension in Matlab 2014a. Correlations were carried out using Age, Gender and total grey matter volume as covariates. The cluster-level threshold was set at $P<0.05$ (FWE-corrected), with a voxel-level threshold set at $P<0.001$ uncorrected. A correction for the non-stationarity of smoothness was also applied using the NS toolbox (http://fmri.wfubmc.edu/cms/software#NS), allowing cluster-level statistics in VBM data. The XJ view toolbox was used to localise significant clusters.

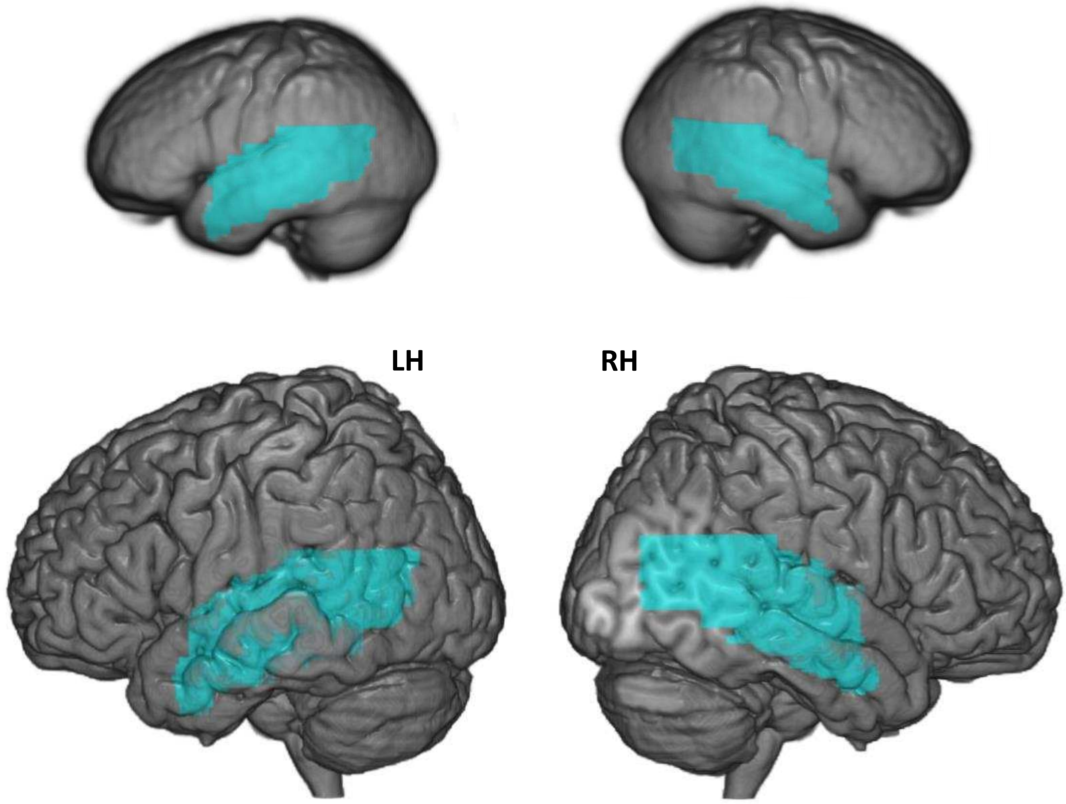


*Figure 3-14 ROI masks covering the STG and MTG, shown on the averaged brain used in the VBM analysis (top) and shown on more detailed brain exemplar (bottom - but less representative of the mask in terms of coordinates).*

### 3.2.2.2.1.1 Explicit AV timing

The location of clusters whose grey matter density correlated significantly with performance in explicit timing tasks are shown in Figures 3-15, 3-16 and 3.17, on the next page. Clusters were derived using the xjview toolbox (http://www.alivelearn.net/xjview) for SPM, Matlab. The ability to discriminate AV temporal order (TOJ SD) correlated negatively with a cluster of grey matter volume in the right hemisphere. Smaller values of the TOJ SD indicate better AV temporal order discrimination, thus individuals who showed higher sensitivity to AV temporal order tend to have larger volumes of grey matter in this cluster. This cluster was situated within the Superior temporal sulcus, and superior and middle temporal gyrus (MTG/STS/STG) (peak coordinates: 58,-12,-7, *p*=.001, FWE corrected at the cluster level). Performance in discriminating between synchronous and asynchronous AV stimuli (SJ SD) correlated negatively with a cluster in right posterior superior and middle temporal gyrus (pSTG/pSTS/pMTG) (peak coordinates: 59, -62, 17, *p*=.032, FWE corrected at the cluster level). The Window of simultaneity (SJ Win) derived from the SJ task correlated negatively with grey matter volume in a cluster located in the right superior temporal gyrus and right middle temporal gyrus (STG/MTG) (peak coordinates: 57,-15-6, *p*=.006, FWE corrected at the cluster level)

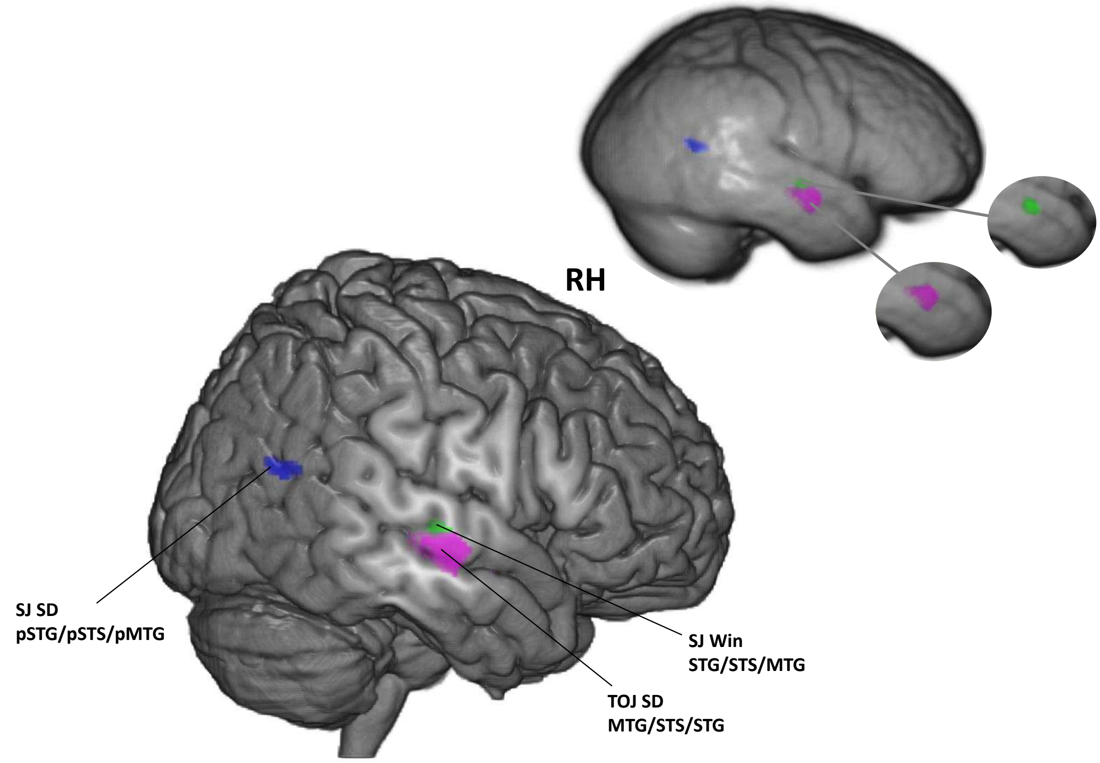


*Figure 3-15: Results of the MTG/STG ROI analysis for explicit timing measures shown (top right) on averaged brain used in analysis (bottom) shown on more detailed brain exemplar (but less representative of exact coordinates of the clusters identified).*

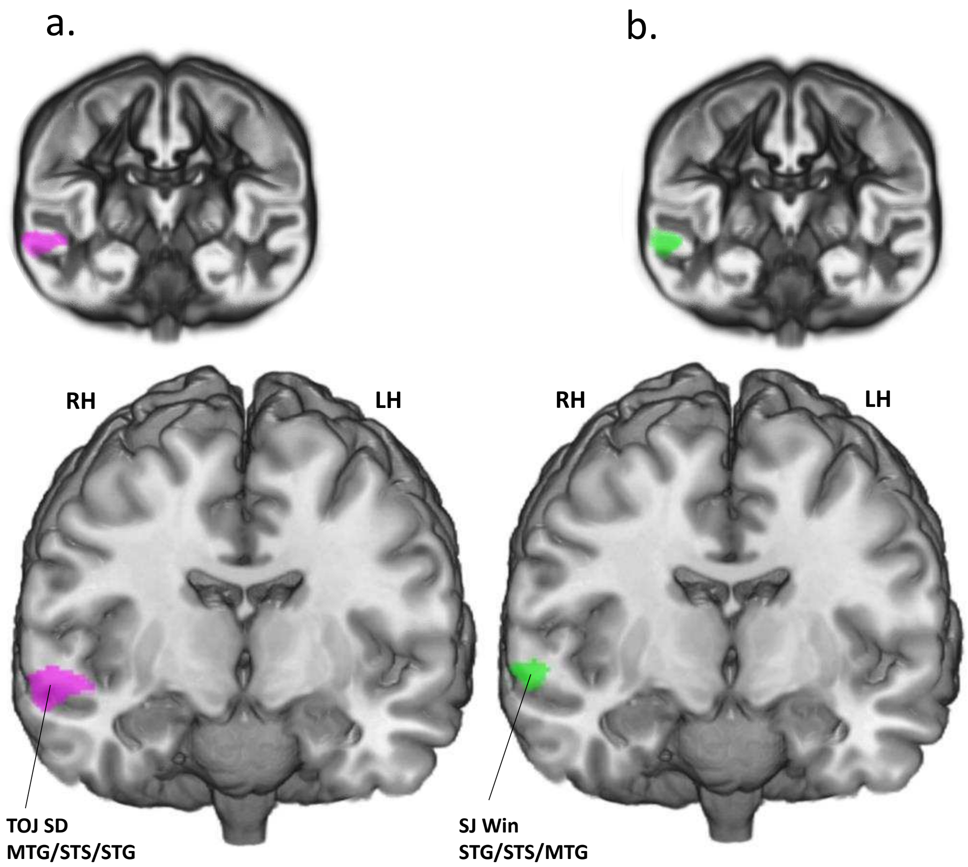


*Figure 3-16 Results of the MTG/STG ROI analysis for **a. TOJ SD** and **b. SJ win** on averaged brain used in analysis (top) and shown on more detailed brain exemplar (but less representative of exact coordinates of the clusters identified) (bottom).*

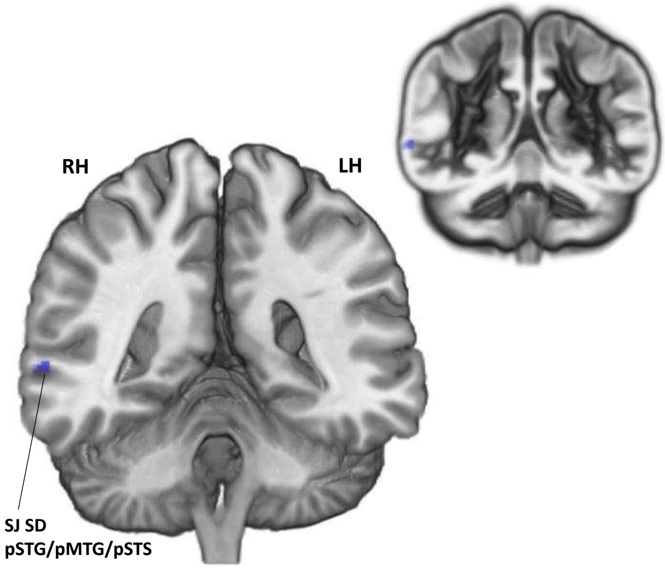


*Figure 3-17: Results of the MTG/STG ROI analysis for **SJ SD** on averaged brain used in analysis (top) and shown on more detailed brain exemplar (but less representative of exact coordinates of the clusters identified) (bottom).*

### *3.2.2.2.1.2 Implicit AV timing and AV integration*

The location of clusters whose grey matter density correlated significantly with higher susceptibility to the McGurk illusion are shown in Figures 3.18, 3.19 and 3.20. Performance in implicit discrimination of AV synchrony (McG SD) did not correlate significantly with grey matter density. The window of AV integration (McG Win) correlated negatively with a cluster of grey matter volume in the right hemisphere, situated within the posterior middle temporal gyrus, (pMTG) (peak coordinates: 48,-72, 14, *p*=.013, FWE corrected at the cluster level). McGurk integration (McG Max) correlated positively with a cluster in superior temporal gyrus (STG) (peak coordinates: 65,-2, -6, *p*=.032, FWE corrected at the cluster level).

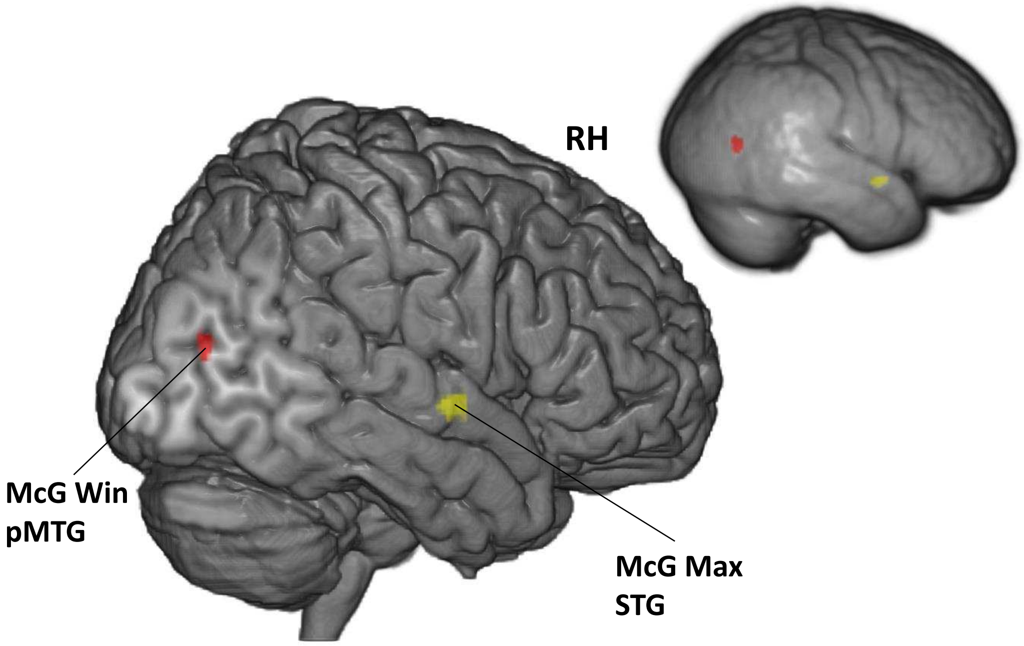


*Figure 3-18: Results of the MTG/STG ROI analysis for* ***McG Win*** *and* ***McG Max*** *displayed on averaged brain used in analysis (top right) and shown on more detailed brain exemplar (but less representative of exact coordinates of the clusters identified) (bottom left).*

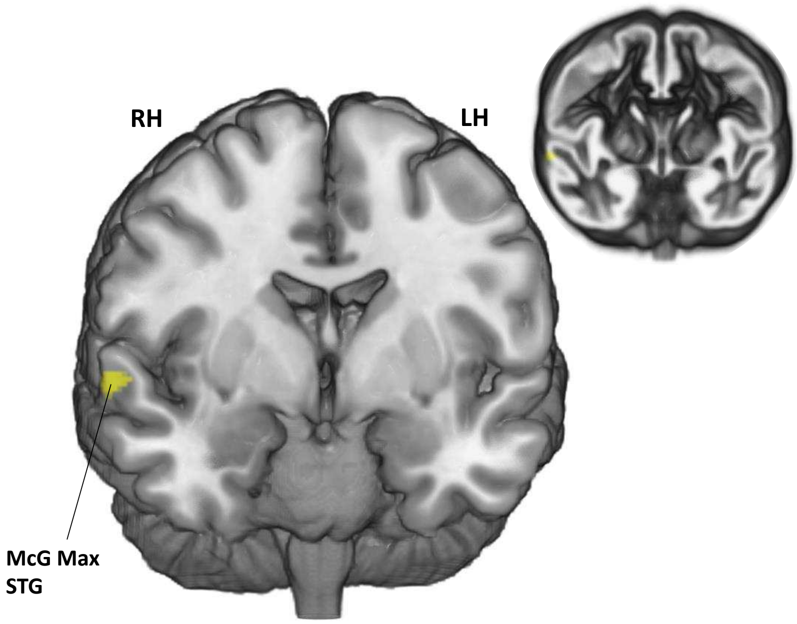


*Figure 3-19: Results of the MTG/STG ROI analysis for **McG Max** displayed on averaged brain used in analysis (top right) and shown on more detailed brain exemplar (but less representative of exact coordinates of the clusters identified) (bottom left).*

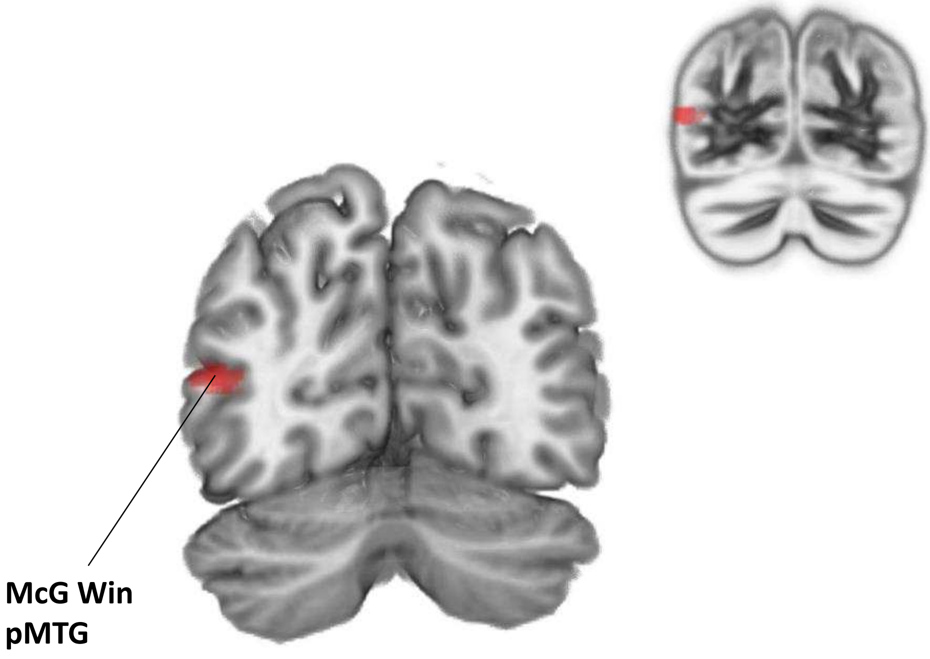


*Figure 3-20: Results of the MTG/STG ROI analysis for **McG Win** displayed on averaged brain used in analysis (top right) and shown on a more detailed brain exemplar (but less representative of exact coordinates of the clusters identified) (bottom left).*

***Cluster Level Statistics: Bilateral Superior/ Middle Temporal Cortex ROI analysis***

| **Measure** | **Contrast** | **Hemi** | **Structure** | **xyz (mm, MNI)** | **$K_{voxel}$** | **$p_{corrected}$** |
|---|---|---|---|---|---|---|
| ***TOJ SD*** | Negative | R | MTG/STS/STG | 58, -12, -7 | 713 | *p*=.001 |
| ***SJ SD*** | Negative | R | pSTG/STS/pMTG | 59, -62, 17 | 94 | *p*=.032 |
| ***SJ Win*** | Negative | R | MTG/STS/STG | 57, -15, -6 | 419 | *p*=.006 |
| ***Integration Win*** | Negative | R | pMTG | 48, -72, 14 | 69 | *p*=.013 |
| ***Maximum Integration*** | Positive | R | STG | 65, -2, -6 | 127 | *p*=.032 |

*Table 3-8: Cluster level statistics temporal cortex MTG/STG ROI analysis. Smaller behavioural measures of TOJ SD, SJ SD, SJ Win and Integration Win represent higher sensitivity to AV asynchrony. Smaller McG Max measures represent less susceptibility to the McGurk illusion*

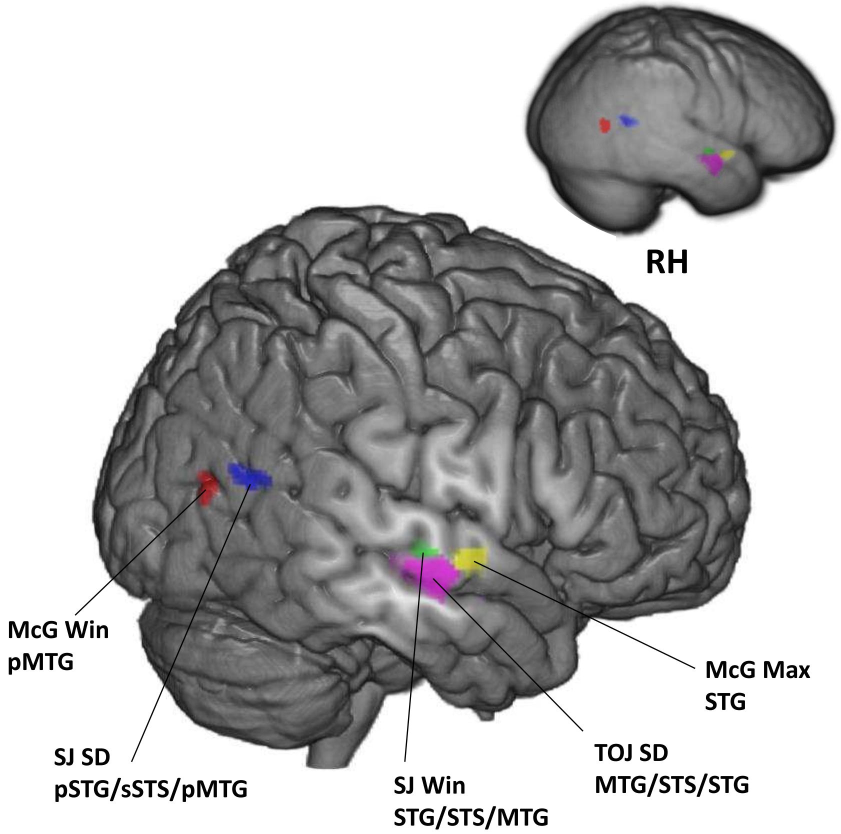


*Figure 3-21: Results from MTG/STG ROI analysis across implicit and explicit tasks on averaged brain used in analysis (top right) and on a more detailed brain exemplar (but less representative of exact coordinates of the clusters identified).*

#### 3.2.2.2.2 Meta-mask and whole brain analyses

A second analysis was performed using a mask containing all the areas reported and cited in the introduction to be involved in AV timing and integration, in order to search for any additional structures that may be related to performance in explicit and implicit temporal processing and AV integration. The masks used for the analysis can be seen below, in Figure 3.22.

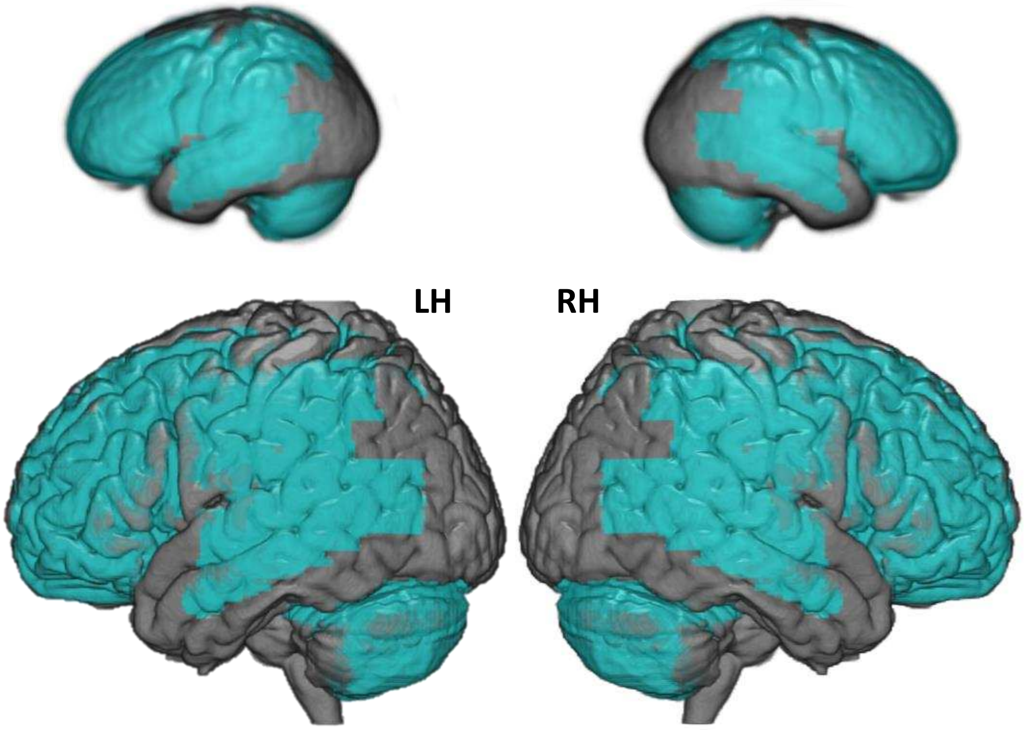


*Figure 3-22: Meta-mask based on previous fMRI literature on temporal order processing, synchrony processing and McGurk AV integration on averaged brain used in analysis (top ) and on a more detailed brain exemplar (but less representative of exact coordinates of the mask).*

The analysis revealed an additional positive relationship approaching significance between McGurk susceptibility (McG Max) and two clusters of grey matter volume, located in the inferior parietal lobe bilaterally (peak coordinates L: -45, -36, 56, *p*=.068, FWE corrected at the cluster level; peak coordinates R: 47, -41, 48, *p*=.068, FWE corrected at the cluster level). These are shown in Figure 3.23, on the next page.

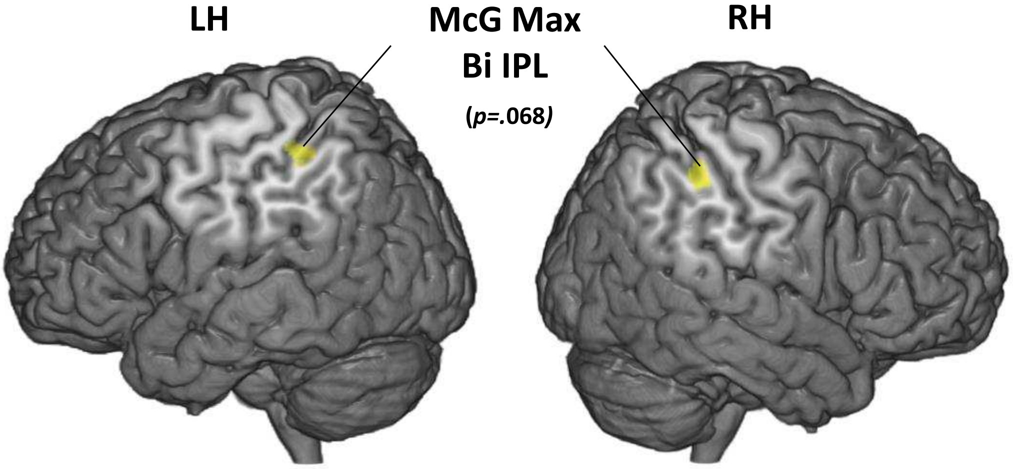


*Figure 3-23: Results of Meta-mask ROI analysis for McG Max (p=.068)*

***Cluster Level Statistics: Meta-Mask ROI analysis***

| **Measure** | **Contrast** | **Hemi** | **Structure** | **xyz (mm, MNI)** | $K_{voxel}$ | $p_{corrected}$ |
|---|---|---|---|---|---|---|
| ***McG Max*** | Positive | L | IPL | -45, -36, 56 | 163 | *p*=.068 |
| ***McG Max*** | Positive | R | IPL | 47, -41, 48 | 125 | *p*=.068 |

*Table 3-9: Cluster Level Statistics of Meta-Mask ROI results. Smaller McG Max measures represent poorer performance.*

### 3.2.2.2.3 Whole brain analysis

A whole brain analysis was also run in order to search for any additional structures that may be related to performance in explicit and implicit temporal processing and AV integration. No significant correlations aside from the previously revealed correlation between TOJ SD and grey matter density were found.

# 3.3 DISCUSSION

### 3.3.1.1 IMPLICIT AND EXPLICIT AUDIOVISUAL TEMPORAL PROCESSING AND BRAIN STRUCTURE

The present findings are the first to demonstrate that the extent to which individuals are able to synchronise and integrate AV information is reflected in structural individual differences in cortical grey matter density. Greater sensitivity to AV asynchrony during explicit temporal judgements, smaller temporal windows of AV synchrony perception and AV integration as well as higher susceptibility to the McGurk illusion were associated with increased grey matter volume in distinct subregions of the right temporal cortex, suggesting that AV relative timing may be processed across multiple, task-specific mechanisms, both within explicit and across implicit and explicit tasks. Strong positive behavioural correlations between measures derived from the explicit temporal judgements, and a failure to find positive correlations between analogous parameters derived from the implicit and all the explicit temporal profiles suggest that the temporal mechanisms underlying SJs and TOJs may work in agreement with one another, and somewhat independently from temporal mechanisms underlying AV integration.

Behavioural measures of AV temporal discrimination ability (SD) across the two explicit AV timing tasks correlated positively, but were related to different clusters of grey matter volumes. Higher sensitivity in discriminating AV synchrony from asynchrony during explicit temporal order judgements (TOJ SD) was related to increased grey matter volume in a cluster located in right MTG/STS /STG. Higher sensitivity in discriminating AV synchrony from asynchrony during explicit synchrony judgements (SJ SD) was related to increased grey matter volume in a cluster also located across MTG/STS/STG, but more posteriorly. SJ SD measures were statistically independent from SD measures derived from the McGurk integration task, and the latter were not significantly correlated with grey matter volume. Thus, at the behavioural level, the degree to which individuals are sensitive to AV asynchrony is consistent across qualitatively different explicit temporal judgements, but seems to be less consistent with the degree to which individuals are implicitly sensitive to AV asynchrony during AV integration, as McG SD correlated with TOJ SD but not with SJ SD, and the relationship between TOJ SD and McG SD much weaker than the relationship found between discrimination ability across the two explicit temporal judgement tasks.

Correlation analysis also revealed that the width of window of AV synchrony (SJ Win) was statistically independent from the width of the temporal widow of AV integration (McG Win). These measures were also distinct at the neural level. Smaller measures of the width of window of synchrony (SJ Win) were related to increased grey matter volume in STG/STS/MTG, a cluster which was close to and overlapping the area related to temporal sensitivity from the TOJ task. Smaller measures of the temporal window of AV integration (McG Win) were however related to increased grey matter volume in posterior MTG. The lack of a positive correlation between these two analogous measures and their distinct structural correlates suggest that AV integration and explicit synchrony processing might rely on distinct underlying temporal mechanisms.

Audiovisual integration is assumed to be contingent upon whether auditory and visual events are seen to belong together. According to this view, whilst the assumption of belongingness can be promoted by AV synchrony, belongingness itself can affect whether AV events are perceived to have occurred at the same time (Spence, 2007; Vatakis, Ghazanfar, et al., 2008; Welch & Warren, 1980). Under this premise, individual ability to discriminate between synchronous and asynchronous AV stimuli should be consistent across AV integration and explicit timing judgements. On the contrary, the null relationships observed here between both temporal parameters extracted from the SJ and AV integration functions suggest that the ability to discriminate AV synchrony across SJs and implicit judgements is not consistent within individuals. Additional analyses revealed that on average, the window of synchrony was significantly larger than the window of AV integration, concurring with previous studies which have compared the two processes (Martin et al., 2012; van Wassenhove et al., 2007). However, the novel finding is that the two were statistically independent.

The SD parameters from the TOJ and SJ fits were strongly and positively correlated, despite being related to distinct structural correlates. There is a possibility that the strong behavioural relationship observed between TOJ and SJ might be attributed to structural and/or functional connectivity. All the parameters derived from the explicit temporal profiles are related to clusters located along the STS/MTG/STG, whereas the parameter extracted from implicit temporal profiles correlates with a cluster in MTG only. Due to this, connectivity between the different neuronal populations related to explicit timing might be stronger compared to the connectivity between implicit and explicit neuronal populations because the former are roughly located within the same cortical areas. This is of course only a conjecture but could be tested using structural imaging methods. For example Diffusor Tensor Imaging analyses could reveal whether structural connectivity between different clusters differs as a function of which parameter they are associated with.

Psychophysical research has shown that TOJs and SJ tend to produce uncorrelated measures of the PSS leading to the argument that these two processes are supported by distinct underlying mechanisms (Love, Petrini, Cheng, & Pollick, 2013; van Eijk et al., 2008; Weiss & Scharlau, 2011). Some have even gone as far as to argue that TOJ are not a valid measure of AV temporal order processing (García-Pérez & Alcalá-Quintana, 2012). The current investigation however suggests that this may not be the case, at least when measuring the ability to discriminate between synchronous and asynchronous AV stimuli. Behaviourally, sensitivity to AV asynchrony in the TOJ task does in fact correlate with ability in performing the SJ task and is also to some degree related to performance in implicit temporal processing when sensitivity measures are not dependent on the PSS and when the types of psychometric functions fitted to the data are based on the same theoretical principles (Yarrow et al., 2011). For instance, the SD measure used here to quantify performance is represented by the standard deviation of a Cumulative Gaussian function for SJ, or the average of two standard deviations of two cumulative Gaussian functions for SJs or AV integration. The standard deviation does not depend on where the PSS happens to lie on the x axis, but on how abrupt the transition is between the two response types as asynchrony is varied. As performance in TOJs correlates with temporal parameters from SJs and with the window of AV integration, TOJs not only can provide an estimate of how good an individual is at discriminating temporal order specifically, but might also give some idea of how that individual might perform certain aspects of other AV temporal processing tasks.

Higher McGurk susceptibility was related to increased grey matter volume in the right STG. Nath and Beauchamp (2012) reported individual differences in susceptibility to the McGurk effect correlated with strength of the BOLD signal in the left STS. Several factors could account for the difference between these findings. Firstly, the current investigation searched for areas which differ in terms of structure between individuals, rather than in terms of function. The relationship between structure and BOLD activity is yet to be clarified and it is uncertain as to whether an area showing higher BOLD signals would necessarily also have more grey matter volume. One study which measured both BOLD activity and grey matter volume in the same participants found that only a proportion of the areas found to contain increased grey matter volume also showed increased BOLD activation (Kim et al., 2010) and also reported greater BOLD activity in areas which did not also show increased grey matter volume. Similarly, Benedetti et al. (2009) measured BOLD responses and grey matter volume in schizophrenia and found areas which showed BOLD activity differences but no structural differences. Altogether, the small number of studies that have measured BOLD responses and grey matter volume in the same participants do not suggest that variability in structure and in BOLD responses are always necessarily linked.

Secondly, Nath and Beauchamp (2012) selectively searched within the left STS rather than searching bilaterally for activation differences. In addition, the measures of McGurk susceptibility correlated to BOLD signal strength in Nath and Beauchamp's study were derived from judgements made on synchronous AV stimuli. As mentioned in the introduction to this chapter, AV synchrony may not necessarily be the optimal AV relative timing for AV integration for all participants, and the asynchrony at which AV integration occurs maximally is subject to wide inter-subject variability. Presenting AV information synchronously to all participants might therefore result in behavioural individual differences related to the ability to integrate AV information but also related to the ability to concurrently synchronise AV information. Therefore the area identified by Nath and Beauchamp could reflect individual differences in both integrating and synchronising AV information. Conversely, the area identified here is likely to reflect only the ability to integrate AV information. This is because susceptibility to the McGurk illusion (McG Max) in this study is not restricted to any particular AV asynchrony, and thus does not depend on whether the optimal asynchrony for AV integration is close to physical synchrony or not.

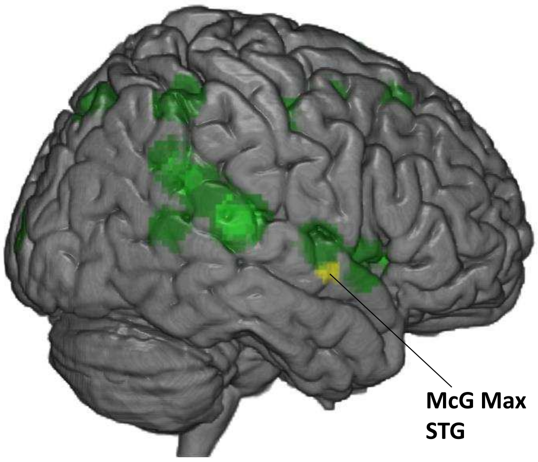


*Figure 3-24: Grey matter volume cluster in right STG correlated with Susceptibility to the McGurk illusion (Yellow) in the context of AV Integration areas previously identified by fMRI studies (green)*

The results of this chapter concur with other previous fMRI literature, as can be seen in Figures 3.24 and 3.25 (latter located on the next page). The superior temporal cortex has been previously implicated in AV temporal processing (Stevenson et al., 2011; Stevenson et al., 2010). Regions of the posterior STS, located within the STC have been implicated in explicit perception and processing of AV synchrony and was reported to be responsive to the temporal order of AV events during synchrony judgements (Noesselt et al., 2012). The STG has been reported to be active during auditory unisensory temporal order judgements (Moser et al., 2009) as well as during unisensory visual synchrony judgements (Lux et al., 2003). The STG was reported by Szycik et al. (2012a) to show increased activation bilaterally during AV integration of incongruent AV stimuli compared to when the stimuli were perceived veridically. Here, an additional positive relationship between grey matter volume in bilateral IPL and individual differences in McGurk susceptibility was found to approach statistical significance. The IPL has been reported to exhibit superadditive responses to AV non-speech stimuli (Calvert et al., 2001), as well as increased activation to stimuli which elicit the McGurk illusion (Jones & Callan, 2003; Skipper et al., 2007).

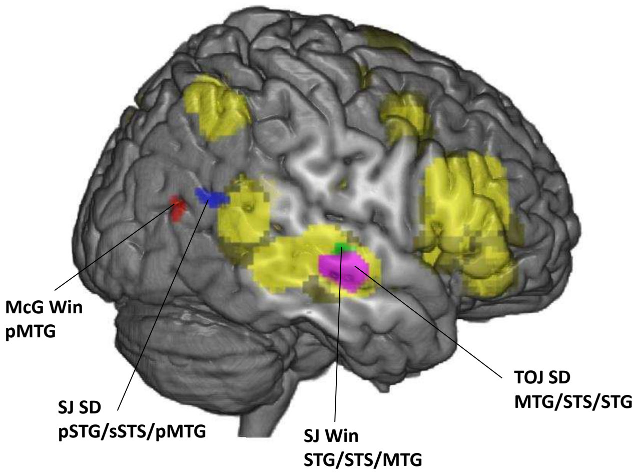


*Figure 3-25: Grey matter volume clusters in right temporal cortex correlated with TOJ SD (Pink), SJ SD (Blue) SJ Win (Green) and McG Win (Red) in the context of areas previously identified as related to AV temporal processing by fMRI studies (yellow).*

The temporal window of AV integration (McG Win) and susceptibility to the McGurk illusion (McG Max) correlated with distinct clusters of grey matter volume in the posterior MTG and anterior STG, respectively. This broadly concurs with fMRI findings reported by Stevenson, vanDerKlok, Pisoni, & James (2011), who identified a subregion bilaterally within the STC showing increased BOLD activity exclusively for AV synchrony and another subregion showing increased activity exclusively for AV integration. Overall, this supports the notion that different complementary processes of AV integration are supported by distinct neural substrates. The current findings extend those of Stevenson et al. by showing that these qualitatively different aspects of AV processing can also be distinguished structurally at the neural level, and that individual differences in performance at these processes are related to structural variability of their neural substrates.

### 3.3.1.2 LATERALITY OF RESULTS

The current structural correlations with temporal processing measures are all located in the right hemisphere. Although the left hemisphere has been argued to also have an advantage over the right hemisphere in terms of temporal processing (Nicholls, Gora, & Stough, 2002; Nicholls, 1994), there is also evidence to suggest against a left hemisphere dominance. Behavioural evidence for hemispheric equivalence (Brown & Sainsbury, 2002) as well as for the notion that the two hemispheres support different types of temporal processing. For example, Okubo & Nicholls (2008) reported that the left hemisphere was dominant when stimuli had a shorter duration, whereas the right hemisphere showed dominance for longer stimulus durations (>240ms). Other research suggests right hemisphere dominance for multisensory temporal processing. Using behavioural methods, Spence, Shore and Klein (2001) found that visual-tactile JNDs were significantly smaller when stimuli were presented on the left visual field (right hemisphere) compared to the right, suggesting a right hemisphere advantage for processing of temporal order. Wittmann and Burtscher (2004) reported that correlations between the size of lesions and auditory JNDs were stronger in the right hemisphere compared to the left. Funnell, Corballis and Gazzaniga (2003) reported that the right hemisphere in split brain patients showed superior performance in visual temporal order judgements compared to the left hemisphere.

Although some fMRI studies have identified areas in the left rather than the right hemisphere to be involved in AV integration (Beauchamp et al., 2004; Calvert, Campbell, & Brammer, 2000), areas in the right hemisphere have also been implicated in AV integration. Benoit et al. (2010) and Sekiyama (2003) reported increased bilateral activation in the STS during integration of AV speech. Stevenson and James (2009b) showed that audiovisual speech in noise detection activated bilateral STS. Baum, Martin, Hamilton, & Beauchamp (2012) reported a patient whose ability to integrate AV speech was not diminished after complete destruction of the left STS and adjacent areas, showing that the ability to integrate AV information is not lost once the left temporal areas are destroyed and that the right hemisphere is also involved in AV integration. The patient exhibited no activity in the lesioned left STS, but compared to healthy controls, the patient showed activation of a larger area of right STS, as well as higher response amplitude in the right STS. The authors also reported bilateral activation of the STS in control participants during processing of McGurk stimuli, showing the AV integration mechanisms are not restricted to the left hemisphere. Although here the ROI also contained the left superior and middle temporal gyri, we found no correlations between grey matter volume and behaviour in the left temporal cortex. One possible explanation is that left hemisphere is as well developed as it can be in most people, and those who perform better than average on temporal processing and AV integration tasks use better developed supplementary resources in other parts of the brain, for example in the corresponding anatomical areas of the right hemisphere.

#### 3.3.1.3 STRUCTURE VERSUS ACTIVITY: POSSIBLE SIMILARITIES AND DIFFERENCES

In comparison to the functional imaging studies reviewed, the number of areas structurally related to variability in AV temporal processing and AV integration processes found in the current investigation is very small. One of the reasons is likely to be related to differences between the way in which behavioural measures were correlated with brain function in fMRI studies and with anatomical structure in the current experiments. For example the current investigation correlated variability in grey matter density with individual differences in very specific aspects of AV timing and integration performance. In contrast, fMRI studies have typically compared functional activation which correlates with exposure to particular types of stimuli or with particular percepts. In terms of AV integration for example, brain activity is averaged over trials in which the McGurk illusion is likely to be perceived or has been perceived and compared against brain activity averaged across trials in which the illusion is unlikely to be perceived or in which it has not been perceived. Exposure to a given stimulus is likely to activate a wide network of areas, some which support processes that are central to the process being investigated and others which are only peripheral. In contrast, grey matter volume here was correlated with a very specific aspect of AV integration. The McGurk susceptibility measure correlated here reflects an individual’s maximum ability to integrate AV information, regardless of the AV asynchrony at which the stimuli is presented. This is because here, AV stimuli were presented to participants at various AV asynchronies, and McGurk susceptibility was estimated from the peak of the temporal profile of AV integration, regardless of its position on the AV asynchrony range.

The structural results here also reflect areas which differ across individuals whereas fMRI results usually reflect areas that respond to stimuli or conditions in the same way across individuals. Brain areas identified in functional studies are likely to support online temporal and integration processes, regardless of how efficient these are and what their perceptual outcome is, whereas the areas identified here are likely to have some contribution to how well these processes perform when they are active. The behavioural measures correlated here and in functional imaging studies therefore represent two qualitatively different things, leading to anatomical correlates which likely support different constructs. Some of the areas identified in the current investigation overlap with areas previously identified in fMRI research to be related to AV processing. It could be argued that the areas identified here support the critical functional role of those in fMRI literature with which they overlap.

Correlations between performance in a given task and larger volumes of grey matter in areas related to its performance are indicative that a given mechanism is better equipped with the resources necessary for performing the neural computations it is responsible for. Grey matter volume measured by MRI however consists of various substructures, ranging from neuronal cell bodies, dendrites, axon terminals to glial cells. Which of these substructures contribute the most to individual differences in behaviour is still unclear (Kanai & Rees, 2011). However, one possibility is that the amount of neurons within a given structure is related to better performance. Roach, Heron, Whitaker, and McGraw (2011) proposed that AV relative timing is estimated from the distribution of activity derived from different neural populations responsive to various AV temporal delays. A distribution of signals produced by larger populations of neurons is likely to be subject to less noise compared to one produced by a small neuronal population. Less noise would lead to more reliable distributions, which might in turn be related to better behavioural discrimination of AV synchrony or AV temporal order.

### 3.3.1.4 CONCLUSIONS

In summary, the current chapter showed that structural differences in cortical grey matter density are related to the extent to which individuals are able to synchronise and integrate AV information. Greater sensitivity to AV asynchrony in implicit and explicit AV timing, smaller temporal windows of AV integration and synchrony perception and higher rates of AV integration were associated with greater density of grey matter volume in distinct subregions of the right temporal cortex. Clusters related to performance in explicit temporal judgement tasks were located along the STS and overlapped with MTG and STG volume whereas the cluster which correlated with the window of AV integration was located more posteriorly, within the MTG. The analysis also revealed that the parameters representative of the window of AV integration and susceptibility to the McGurk illusion were related to distinct areas of the right temporal cortex, suggesting that complementary processes of AV integration might be supported by different neural substrates and showing that the performance of these processes is reflected in structural brain differences. Behaviourally performance across the two explicit tasks correlated positively, whereas performance across implicit and explicit tasks was inconsistent. Altogether, these results indicate that AV relative timing across implicit and explicit AV temporal judgements is likely to be processed via multiple task-specific mechanisms, supported by distinct neuronal populations.

# 4 CHAPTER 4: READ MY LIPS. AUDIOVISUAL TIMING AND INTEGRATION AND DYSLEXIA

## 4.1 INTRODUCTION

Associations between visual and auditory speech begin to form early on in development, prior to learning to read. For instance, even before they learn to speak, infants are susceptible to the McGurk effect (Kushnerenko, Teinonen, Volein, & Csibra, 2008; Rosenblum, Schmuckler, & Johnson, 1997), which requires some knowledge of the correspondence between speech sounds and the lip-movements which produce them. Forming multisensory associations between auditory and visual speech is thought to contribute to the development of language (Teinonen et al., 2008) and may be a prerequisite of learning grapheme-phoneme correspondences (Blomert & Froyen, 2010). The ability of children as young as 7 months to integrate audiovisual (AV) speech (Hollich et al., 2005) as well as to detect mismatches in previously learned AV pairs (Gogate & Bahrick, 1998) has been shown to be affected by the relative timing of auditory and visual speech information. This indicates that even at an early, pre-linguistic stage of development, processing AV relative timing is important in AV integration. Deficits in temporal processing could therefore affect the ability to form AV correspondences which later may be useful in learning to pair graphemes and phonemes automatically, leading to reading impairments later in life.

Dyslexia for example, is a neurobiological condition characterised by problems with word recognition, spelling and decoding, despite otherwise typical reading instruction and educational or professional attainment (Lyon et al., 2003). Individuals diagnosed with this condition show reduced automaticity in grapheme-phoneme association which is likely to be a result of poor learning of grapheme-phoneme correspondence (Ramus, 2001). In turn, this could be partly the result of deficits in AV temporal processing and/or integration. Only a limited amount of research is however available on the relationship between AV processing and typical reading as well as on the ability of dyslexic individuals to integrate and synchronise AV information. The current chapter addresses this gap in the literature.

The following chapter will review research on unisensory speech and non-speech temporal and sensory processing in dyslexia and in typical readers. This section will be followed by a discussion of what is so far known about crossmodal temporal processing and integration in relation to typical and dyslexic reading. The four experiments in this chapter compare performance of dyslexic and typical readers in different aspects of implicit and explicit AV temporal processing and in AV integration, using both speech and non-speech stimuli. The relationship between reading ability across the whole sample and performance in different aspects of AV integration and timing is also explored.

### 4.1.1 PHONOLOGICAL ACCOUNTS OF DYSLEXIA

Phonological processing accounts of dyslexia define the disorder as a specific language impairment and argue that reading impairments in the disorder are caused exclusively by left hemisphere perisylvian dysfunction which leads to deficient phonological representations, in turn affecting grapheme-phoneme mapping and other phonological skills (Goswami, 2000; Stanovich, 1988). This explanation however fails to account for the many other deficits that are associated with dyslexia, such as poorer visual, auditory (Hämäläinen, Salminen, & Leppänen, 2013; Laasonen, Service, & Virsu, 2001) and temporal processing (Farmer & Klein, 1993; Farmer & Klein, 1995). Such deficits, which are not phonological in nature, suggest that phonological processing impairments characteristic of dyslexia may stem from lower level sensory processing problems (Farmer & Klein, 1993; 1995). The evidence for this mainly comprises studies that have used visual and auditory speech and non-speech stimuli in isolation and less focus has been placed on whether the integration or temporal processing of multisensory information is affected in the disorder.

### 4.1.2 SPEECH PROCESSING: AUDITORY AND VISUAL

Dyslexic individuals have been reported to show poorer performance in interpreting auditory speech-in-noise (SPN) (Ramirez & Mann, 2005; Savage et al., 2005), both when noise is added to speech externally and when noise is introduced by degrading the speech itself (Ziegler, Pech-Georgel, George & Lorenzi, 2009). Ziegler et al.  reported that performance in SPN detection tasks was a predictor of reading ability in dyslexia even when memory, attention and low-level sensory processing were controlled for, suggesting that at least a proportion of reading impairments in dyslexia may be accounted for by difficulties in speech processing mechanisms. Other studies did not however find the same association of deficits (Hazan, Messaoud-Galusi, Rosen, Nouwens, & Shakespeare, 2009), possibly as a result of variations of stimulus types, and task difficulty (Ramus, 2003).

Visual-only lip reading ability is also poorer in dyslexic children (Bastien-Toniazzo et al., 2010; Campbell, Whittingham, Frith, Massaro, & Cohen, 1997) and adults in whom it has also been reported to correlate with reading ability (Mohammed, Campbell, Macsweeney, Barry, & Coleman 2006). Altogether, the findings suggest that impairments in dyslexia may not be restricted to orthographic representations of language, but that they also generalise to spoken language.

### 4.1.3 UNISENSORY TEMPORAL PROCESSING IN DYSLEXIA

#### 4.1.3.1 AUDITORY TEMPORAL PROCESSING

Some evidence suggests that speech processing deficits in dyslexia (such as the those mentioned above), result from impaired processing of temporal information (Hornickel, Skoe, Nicol, Zecker, & Kraus, 2009; Schulte-Körne, Deimel, Bartling, & Remschmidt, 1999, 2001). For example, dyslexic individuals have poorer sensitivity to the temporal order of speech sounds, which improves once consonant pairs are lengthened in duration (Rey, De Martino, Espesser, & Habib, 2002), but not if consonant complexity is reduced, or the inter-stimulus duration is lengthened by inserting a neutral vowel within consonant clusters (De Martino, Espesser, Rey, & Habib, 1999). This suggests that impairments in processing successive consonant sounds may stem from difficulties in processing and categorising short sounds, rather that processing successiveness. In order to test this hypothesis, vandermosten et al. (2010, 2011) examined auditory categorisation ability in dyslexic children and adults. Speech and non-speech stimuli were varied either in terms of temporal cues or of non-temporal cues, and acoustic complexity was controlled for across all the stimuli categories. Dyslexic participants in both studies performed as well as controls on categorisation of sounds that differed based on non-temporal cues, but showed poorer ability in categorising sounds that differed purely by their temporal cues, regardless of whether they were speech or non-speech.

Impairments in processing temporal properties of speech in dyslexic children have also been demonstrated at the neural level. Meng et al. (2005) measured mismatch negativity (MMN) responses to deviations in temporal and spectral properties of auditory speech and non-speech stimuli, using EEG in Chinese children. The MMN response is an event related potential (ERP) component which is elicited by presenting an odd stimulus in a repetitive sequence of stimuli. Abnormalities in this component are thought to represent attentional or sensory processing impairments (Garrido, Kilner, Stephan, & Friston, 2009). In the speech conditions, MMNs in response to deviations in the initial consonant of syllables such as ‘ga’ and ‘da’, as well as to deviations in the lexical tone of the stimuli were smaller in the dyslexia group compared to controls, indicating impairment in detection of both temporal and tonal changes in speech. In the non-speech conditions, MMN responses were smaller in dyslexic children compared to controls for deviations in the duration which separated three tones making up composite tone patterns. Dyslexic MMN responses were however comparable to controls when tones deviated in terms of frequency. These findings demonstrate that the detection of temporal and tonal changes in speech is impaired at the neural level in dyslexia, and that the temporal processing deficit generalises to non-speech auditory processing.

Whether auditory processing deficits in dyslexia are restricted to speech, or stem from basic auditory processing impairments is a debated issue (Farmer & Klein, 1995; Tallal, Miller, & Fitch, 1993; Tallal, 2004). For example, Schulte-körne, Deimel, Bartling, & Remschmidt (1998) argued that auditory processing impairments in dyslexia are specific to speech sounds. Schulte- körne and colleagues based this argument on the observation that MMN responses were weaker in dyslexic children and adults only when they were exposed to deviations in speech sounds such as the syllable ‘da’ embedded within a repetitive sequence of ‘ba’, and not to deviations in the frequency of non-speech sine wave tones (Schulte-körne et al., 1998; Schulte-Körne et al., 2001). Meng et al. however did demonstrate that dyslexic children showed differential responses to deviations in duration rather than in frequency of non-speech sounds.

Poorer temporal processing for non-speech stimuli in dyslexia has also been reported by behavioural research (Boets, Wouters, van Wieringen, & Ghesquière, 2007; Breznitz & Meyler, 2003). Some of the earliest work on low-level auditory processing and its relationship to reading impairment was pioneered by Tallal and colleagues. Tallal (1980a, 1980b) argued that phonological impairments in dyslexia stem from deficits in processing rapidly occurring sounds, early on in sensory processing. Tallal (1980a, 1980b) found that dyslexic performance in tasks requiring processing of brief and successive temporal information, such as auditory TOJs was impaired and correlated with poorer performance in non-word reading ability. Similar findings were reported by Reed (1989), who used temporal order judgements (TOJ) to compare sensitivity to the order of sounds across dyslexic and typically developing children. Reed’s dyslexic group performed worse in judging temporal order compared to controls even when stimuli were separated by as much as 400ms, and regardless of whether the stimuli consisted of brief tone or stop consonant pairs. Dyslexic performance was however comparable to that of controls when judging the temporal order of vowel pairs. Vowel sounds tend to occur less rapidly compared to consonants, suggesting that the impairment is specific to rapidly occurring stimuli. These deficits have also been shown to persist into adulthood and to correlate with reading ability (Kinsbourne, Rufo, Gamzu, Palmer, & Berliner, 1991).

Temporal processing of auditory information seems to be related to reading in typically developing individuals as well. Meng et al. (2005) reported that in their sample of typically developing children performance in TOJs, gap detection and frequency discrimination of non-speech stimuli was found to explain 32% of the variance in phonological awareness. Phonological awareness is an individual’s awareness about the sound structure of spoken words, characterised by skills such as the ability to isolate speech sounds from one another, segment words into their individual phonemes and match phonemes across different words. Performance in these TOJs, gap detection and frequency discrimination was also related to reading fluency when phonological awareness performance was partialled out, suggesting that non-speech auditory temporal processing may be related to multiple, independent aspects of reading.

The evidence for an auditory processing deficit in dyslexia is however fairly inconsistent. Firstly, some studies have not found any differences in terms of temporal or rapid auditory processing in dyslexic participants or poor readers. For example Nittrouer (1999) compared temporal processing ability across good and poor phonologically skilled children, on tasks employing both speech and non-speech stimuli. Nittrouer’s tasks were similar to those used by Tallal and her colleagues (1980), but despite this, Nittrouer did not find any temporal processing differences between poor and good readers. It is worth noting however that Nittrouer's ‘poor reader’ sample were classified as poor readers using the Reading subtest of the Wide Range Achievement Test-Revised and not via a formal diagnosis of dyslexia. Furthermore, poor readers made up only 15% of the 110 children tested. Although Nittrouer argues that this is what would be expected in the wider population as a whole, the small proportion of poor readers may account for why Tallal’s findings were not replicated. Finally, Nittrouer did run correlations between individual differences in reading scores and performance in temporal processing tasks within normal readers, poor readers or across the entire sample. As a result, it is unclear whether the temporal processing problems are likely to have a causal relationship to reading impairment, or whether the two are independent.

Auditory processing deficits in dyslexia are not always related to the temporal features of sounds (Amitay, Ahissar, & Nelken, 2002). Even when temporal or rapid auditory processing deficits have been found, they have been reported to be present in only a subset of reading impaired individuals (Farmer & Klein, 1993) suggesting that these might not be a cause of dyslexia, but more likely a co-morbid impairment. Others have argued that low-level temporal processing deficits are unrelated to speech processing problems that contribute to reading impairment (Rosen & Manganari, 2001). For example, even when poor readers do on average show lower performance in auditory temporal processing tasks, individual differences in this performance is not related to individual differences in reading ability (Rosen, 2003), challenging the notion of a connection between low-level auditory processing and reading impairments.

One possible explanation for the inconsistency in findings is that, if auditory temporal processing were a causal factor in reading ability, dyslexic readers may simply be at the lower end of a continuum on ability in both auditory processing and reading performance. If this were the case, sampling differences across studies, in terms of reading ability of dyslexics and controls could account for situations in which no impairment has been found in dyslexia and in which only a subset of dyslexics show the impairment.

#### 4.1.3.2 VISUAL TEMPORAL PROCESSING

Processing dynamic visual stimuli also seems to be impaired in dyslexia. The magnocellular theory of dyslexia (Stein & Walsh, 1997) argues that impairments in the function of the magnocellular system, such as contrast and coherent motion processing are the root of many of the reading problems that characterise the disorder. The magnocellular system is a subdivision of the geniculostriate visual pathway, which is primarily responsible for timing visual events, motion sensitivity and stable binocular fixation, all of which are said to be essential for normal reading development (Stein, 2001). Dyslexic children and adults are on average less sensitive to low-level dynamic visual stimuli (Chase & Jenner, 1993; Stein & Talcott, 1999) as well as to coherent motion in paradigms employing random dot kinematograms (RDK) (Talcott, Hansen, Assoku, & Stein, 2000; Witton et al., 1998). Talcott et al. (2000) reported that dyslexic, but not control performance improved as a result of increasing dot density, which represented an increase in motion energy, suggesting that lower motion sensitivity in dyslexia may stem from deficits in the signal-to-noise sensitivity of magno cells.

In typical developing individuals, the perceived duration of a visual stimulus is normally reduced after being exposed to a period of high frequency invisible visual flicker of 60hz, a frequency to which magno cells are tuned. Johnston et al. (2008) found that this adaptation effect was absent in a dyslexic sample, providing more evidence for a magnocellular function deficit in dyslexia. The lack of adaptation could not be attributed to attentional factors, since the flicker of the stimuli was not detectable. Adaptation effects on the perception of duration as a result of flickering at lower frequencies, such as 20Hz, which do not affect magno cells, were found to be on average equal in dyslexics and controls, indicating that the adaptation impairment found on average to be present in dyslexic individuals, found was specific to perceptual mechanisms supported by the magnocellular pathway. A subset of Johnston et al.'s dyslexic individuals, who scored particularly low at phonological and literacy tasks also exhibited unusual shifts in perceived duration as a result of adaptation to a lower frequency flicker. The parvocellular system is sensitive to this lower frequency flicker, suggesting that in these particular individuals, this system may also function suboptimally. This finding emphasises the need to differentiate between different subtypes of dyslexia, as the underlying cause of reading problems in this population may not be consistent across different subsets.

As with most theories of dyslexia, the magnocellular account is also under debate. Some studies report no magnocellular impairments despite the presence of dyslexia (Kronbichler, Hutzler, & Wimmer, 2002; Tsermentseli, O'Brien, & Spencer, 2008). Variation in stimulus intensity, shown to modulate dyslexic performance in coherent motion detection tasks (Talcott et al., 2000), is likely to account for some of the inconsistency in findings across different studies employing tasks which measure magnocellular function. Moreover, not all visual processing deficits reported to be present in dyslexia, for example visual attention and perceptual memory impairments (Amitay, Ben-Yehudah, Banai, & Ahissar, 2002) can be directly related to the magnocellular system (Johnston et al., 2008; Ramus, 2003), suggesting that dyslexia cannot be characterised only by magnocellular dysfunction. Some of the magnocellular impairments found are difficult to link to the process of reading. For example, dyslexic individuals have been reported to show poorer contrast sensitivity at low luminance levels (Cornelissen, Richardson, Mason, Fowler, & Stein, 1995) and as reading is generally performed under well-lit conditions, this deficit does not easily account for reading impairments. Thus, greater clarification is needed regarding which aspects of magnocellular function are likely to lead to reading impairment and which are not.

The same issue facing the rapid auditory accounts of dyslexia applies to magnocellular accounts. In several studies demonstrating magnocellular deficits, only a proportion of dyslexic individuals were affected. In some studies, these proportions were as small as 29% (Ramus, 2003). Furthermore, there is also the problem of causality. One study (Olulade, Napoliello, & Eden, 2013) found evidence that magnocellular dysfunction may not be a cause of dyslexia, but rather a consequence of poor reading ability. Neural activity in V5/MT measured during motion processing was found to be weaker in dyslexics compared to age matched controls, but equal to that of reading matched younger controls. The authors also reported that neural activity in the dyslexic group increased as a function of reading ability improvement, after subjects took part in a phonological-based reading intervention.

#### 4.1.3.3 GENERAL TEMPORAL PROCESSING – WITHIN BOTH THE VISUAL AND AUDITORY MODALITIES

Although much of the existing research on sensory processing in dyslexia has investigated temporal processing within single unisensory modalities, there is also some evidence for the argument that such deficits generalise to multiple senses within the same individuals (Farmer & Klein, 1995). The General Temporal Processing deficit theory of dyslexia (Nicolson & Fawcett, 1993; van Ingelghem et al., 2001) proposes that temporal processing deficits within and across both auditory and visual modalities is a causal factor in reading impairments in the disorder.

For example Laasonen et al. (2001) compared dyslexic adolescents to matched controls on temporal order and synchrony judgements within the visual, auditory and tactile modalities. Although there was a large degree of overlap between temporal processing abilities across the two groups in this study, on average, dyslexic readers performed worse than controls in all six temporal acuity tasks. The authors also found that performance on these tasks across the entire sample correlated with different aspects of reading ability such as phonological synthesis and non-word span. van Ingelghem et al. (2001) also tested the hypothesis using auditory gap detection and visual double flash detection tasks. In gap detection tasks participants are presented with trains of white noise, within which they are required to detect short gaps of silence. The visual equivalent of this is the visual double flash task, in which participants have to detect whether a visual stimulus flashed once or twice, testing the ability to detect brief visual inter-stimulus intervals. The authors reported that 70% of their dyslexic sample had significantly higher visual and auditory thresholds, and thus poorer performance, compared to age matched controls. Although van Ingelghem et al., (2001) refer to this temporal deficit as ‘crossmodal’ because temporal processing is affected within more than one modality, there are in fact relatively few studies that have investigated crossmodal temporal processing in dyslexia (but see Marja Laasonen, Service, & Virsu, 2002 in section 4.1.4.2). This gap in literature is addressed in the current chapter. Van Ingelghem et al. (2001) also found, as with visual or auditory-only temporal processing deficits, that not all dyslexic individuals exhibit temporal processing impairments within multiple modalities.

## 4.1.4 AUDIOVISUAL PROCESSING, DYSLEXIA AND READING

### 4.1.4.1 AUDIOVISUAL INTEGRATION

Assuming that unisensory processing problems are present in dyslexia, the question arises as to whether multisensory processing is also affected in the disorder, or whether it remains intact despite unisensory deficits. As mentioned at the beginning of this chapter, strong grapheme-phoneme associations might depend on the ability to integrate AV information during the development of reading. Intuitively, one would predict that unisensory deficits would lead to problems with multisensory processing. However Bastien-Toniazzo et al., (2010) compared AV speech integration in dyslexic children and aged matched controls and found no group differences in AV integration. The authors measured AV speech integration using synchronously presented McGurk stimuli and also measured silent lip reading ability and auditory speech perception across a range of noise conditions. The only task at which dyslexic children underperformed was the silent lip reading task, but despite this, the degree of visual influence on auditory perception in the McGurk conditions did not differ from that of controls. Given that silent lip reading ability correlates positively with the degree to which visual information affects auditory perception (Summerfield, 1992), dyslexic individuals would be expected to also show lower rates of AV integration alongside the poorer silent lip reading ability reported by Bastien-Toniazzo et al. (2010).

Campbell et al. (1997) found similar results. The authors measured AV integration using combinations of a range of auditory speech syllables and lip-movements which normally elicit the McGurk effect, such as ‘ba’, ‘va’, ‘tha’, ‘da’ and ‘ga’, and measured unisensory processing by presenting the stimuli unimodally. The task required participants to report the phoneme they heard in the auditory and AV conditions, and what they believed was uttered in silent speech condition. Dyslexic participants performed significantly worse than controls unimodally, and also showed lower AV integration rates which were interpreted to reflect poor unisensory performance. The authors concluded that unisensory speech processing is impaired in dyslexia, but that multisensory integration processes are intact.

Ramirez & Mann (2005) reported that dyslexic individuals gained less benefit from visual information during speech-in-noise detection, compared to controls and participants with auditory neuropathy, which was also attributed to poorer unisensory processing. In this study, AV speech consisted of congruent consonant-vowel combinations, of which the auditory component was embedded in noise and had to be identified. Dyslexic performance in AV speech-in-noise detection declined as a function of an increase in noise level, more so compared to controls. Silent lip-reading ability measured in the visual-only condition was also poorer in the dyslexic group compared to the other two groups.

In all of the aforementioned AV integration literature, AV stimuli were only presented synchronously. As discussed earlier, dyslexic individuals, even if only a proportion, seem to be affected by temporal processing problems in either the auditory or visual modality, or both. If unisensory temporal processing is impaired in dyslexia, then one may expect to find unusual temporal processing between modalities too, which may in turn affect acuity for perceiving the relative timing of auditory and visual streams. Differences between dyslexic individuals and controls in terms of AV integration may therefore not be apparent when the information is presented synchronously, but they may differ when AV information is presented asynchronously.

Hairston, Burdette, Flowers, Wood and Wallace, (2005) reported that AV integration occurred at larger AV asynchronies in dyslexia compared to typically developing individuals. Hairston et al. used temporal ventriloquism to measure AV integration over a range of AV asynchronies, where the auditory stimulus always lagged the visual. Temporal ventriloquism ( Bertelson, 2003; Morein-Zamir, Soto-Faraco, & Kingstone, 2003) is the influence of the timing of a sound on that of a visual stimuli whereby the former ‘pulls’ the perception of the latter towards it.

Thus, the addition of a sound after a flash leads to the flash being perceived as having occurred later in time than it was presented. This phenomenon is however constrained by AV asynchrony in typical individuals, and Hairston et al. tested whether temporal constraints of this illusion were comparable across dyslexics and typical readers. Participants were required to judge the temporal order of pairs of flashes, with sound and without sound present. In the AV condition, the first flash was presented synchronously with a sound and the second flash was either presented synchronously with the sound followed by the sound by a period ranging from 50 to 350ms.

Hairston et al. obtained the asynchronies at which participants gained benefit from auditory information by carrying out multiple comparisons on the raw data, rather than fitting psychometric function to the data. Measures of visual temporal order sensitivity were compared between the condition in which no sound was presented and each of the AV conditions. Overall, performance was improved by adding sound across both groups, but dyslexic individuals gained significantly more benefit from the auditory stimulus when it was present synchronously with the flashes, as well as over a wider range of AV asynchronies, suggesting that dyslexic individuals have a wider window of integration.

Hairston et al. did not measure AV integration over asynchronies where the auditory information leads the visual, therefore it is unclear whether the window of AV integration is in fact larger for dyslexics, or the simply the same size but shifted along an AV asynchrony continuum. The authors did not examine measures of the asynchrony at which AV integration is optimal either (i.e. at which the most benefit from auditory information is obtained), which might have informed whether the temporal profile of AV integration was indeed wider in dyslexia or whether it was shifted towards auditory-lagging asynchronies. Finally, although reading ability was reported to correlate with visual temporal order sensitivity in the no-sound condition, the authors did not report any correlations between reading ability and individual differences in the degree which individuals benefited from auditory information overall or with the degree to which they integrated AV information at wider AV synchronies, thus it is unclear as to whether the crossmodal benefit observed overall and at wider asynchronies is related to, or co-occurs independently from reading impairments in dyslexia.

The role of AV processing in reading is not well understood and conclusions regarding AV processing and its contributions to reading ability are mixed. Nevertheless, AV training has been reported to have remediating effects on reading ability of dyslexic children, using speech (Lyytinen, Ronimus, Alanko, Poikkeus, & Taanila, 2007; Magnan, Ecalle, Veuillet, & Collet, 2004; Veuillet, Magnan, Ecalle, Thai-van, & Collet, 2007) and non-speech stimuli (Kujala et al., 2001). The underlying mechanisms through which AV training might improve reading ability are not understood, and merit further investigation. Functional MRI research shows that activity of cortical areas implicated in the use of grapheme-phoneme associations (Jobard, Crivello, & Tzourio-Mazoyer, 2003) as well as AV integration and timing, such as the Superior temporal cortex (Nath & Beauchamp, 2012; Noesselt et al., 2012; Stevenson et al., 2011) is reduced in dyslexic children (Blau et al., 2010) as well as adults (Blau et al., 2009) during the perception of speech sound and letter pairs. It is possible that AV training improves the function of these areas and that the effects generalise to reading related processes.

EEG research also points towards an AV processing deficit in dyslexia. Froyen, Willems, & Blomert (2011) measured MMN responses of dyslexic children to speech sounds presented concurrently with letter representations. This MMN response is a component of an auditory ERP, but its amplitude typically increases when visual information is presented at the same time as the auditory. This amplitude increase also occurs relatively early after stimulus onset and thus is interpreted to reflect earlier AV integration (Froyen, van Atteveldt, & Blomert, 2010; Froyen et al., 2008). Froyen et al. (2011) found that the response modulation was not present in dyslexic children. This indicates a deficit in the automatic integration of auditory and visual information early on in sensory processing, which may be a causal factor in impaired reading development.

More direct evidence for Froyen’s claims of early AV integration impairments as a causal factor in reading problems in dyslexia comes from Widmann, Schröger, Tervaniemi, Pakarinen and Kujala's (2012) findings. The authors used non-speech AV stimuli to investigate dyslexic ability in matching visual symbol patterns to auditory sound patterns. Behaviourally, dyslexic children were significantly poorer at performing congruency judgements on the AV stimuli. ERP responses typically seen in normal readers when presented with incongruent AV stimuli occurred significantly later in the dyslexic group compared to controls, after stimulus onset. In the left hemisphere, the amplitude of this response was also smaller by nearly a half in dyslexics compared to controls. The authors also found that the amplitude of the response correlated significantly with reading ability. Another ERP component, which usually occurs later and is associated with behavioural relevance of a stimulus, was present in controls, but not in dyslexic children. Furthermore, gamma band oscillatory responses which are normally associated with synchronisation of neural activity related to the process of binding AV information were also not present in dyslexic group. Altogether, these findings suggest that dyslexic individuals suffer from impairments in AV processing at the behavioural and neural level and that these impairments are related to reading performance.

To summarise, electrophysiological studies point towards an AV integration deficit in dyslexia indicated by different patterns of brain activity observed in dyslexic individuals. On the other hand, the majority of behavioural studies which have investigated AV integration in dyslexia using speech in noise detection paradigms and the McGurk effect argue that AV integration is intact in dyslexia and that differences found in AV integration in this group can be accounted for by unisensory processing impairments. These studies have however presented AV information synchronously, and there is some evidence that the temporal profile of AV integration in dyslexia might be different. As shown in Chapter 2 and 3, there are wide individual differences in the AV asynchrony at which AV integration occurs maximally. Paradigms which present AV information synchronously do not take this into account, and any differences which exist as a result of impaired implicit AV temporal processing may not be clear when AV integration is measured only at synchrony. Differences found in brain activity between dyslexic groups and typical readers might therefore in part reflect AV synchronisation differences present during AV integration, to which the behavioural paradigms used in the aforementioned behavioural literature are not sensitive.

#### 4.1.4.2 EXPLICIT AUDIOVISUAL TEMPORAL PROCESSING

It is possible that abnormalities in AV processing in dyslexia might be attributed to difficulties in temporal processing across modalities, rather than reflect a purely AV integration deficit. Explicit temporal processing of multisensory stimuli is somewhat poorer and correlates with phonological awareness in dyslexic individuals (Laasonen et al., 2002). The authors reported significant differences in temporal order acuity in visuotactile conditions, but only found trends of impaired temporal order acuity in the AV condition. In synchrony judgement tasks, visuotactile and audiotactile temporal acuity was significantly poorer in dyslexics compared to controls, but AV acuity did not significantly differ. Laasonen et al. also reported that within controls, temporal acuity was related to phonological synthesis. Performance at audiotactile and visuotactile synchrony judgements has also reported to deteriorate more with age in dyslexia compared to typical readers (Virsu et al., 2003). Average measures of the point of subjective synchrony were not however reported nor were they taken into account when comparing performance. Temporal acuity thresholds were therefore dependent on physical synchrony and not subjective synchrony. As demonstrated in Chapter 2, subjective synchrony varies widely between individuals and this variability may have contributed towards the null effect.

To summarize, the small amount of research that has investigated AV processing in dyslexia has measured AV temporal processing skills in isolation from AV integration skills. Studies which have found no difference between dyslexic individuals and typical readers in terms of AV integration have therefore not taken into account the that AV asynchrony might need to be artificially adjusted in order to obtain representative measures of AV integration, for both controls and dyslexics. Explicit AV temporal processing has only been investigated by two studies using artificial stimuli comprising brief flashes and beeps, and has not been investigated using speech stimuli. Although AV integration in dyslexia has been investigated as a function of AV synchrony in one study (Hairston et al., 2005), the findings do not reflect the entire temporal profile of AV integration, nor do they inform towards potential relationships between reading ability and AV integration. Whether altered profiles of AV integration also exist for AV speech is unknown. The majority of EEG research into AV processes in AV integration has also predominantly used children, thus making it unclear whether differences found are persistent into adulthood or whether development of AV processing is simply delayed in dyslexia.

The current study therefore examines performance in AV integration as a function of AV synchrony in dyslexic and typical readers, for both speech and non-speech stimuli, using the McGurk and Stream-Bounce illusions. Performance in explicit AV temporal judgements for speech and non-speech is also measured. AV temporal processing ability in the context of both AV integration and explicit temporal judgements is therefore measured and compared between groups. Performance across all the tasks is also correlated with different aspects of reading ability, in order to determine whether any differences found between controls and the dyslexic group can account for reading impairment over and above dyslexia.

# 4.2 EXPERIMENTS 4.1 - 4.4

## 4.2.1 METHODS

### 4.2.1.1 SUBJECTS

Participants comprised of 20 individuals formally diagnosed with dyslexia (11 Females, Mean age=22.06) and 22 control participants (15 Females, $M_{Age}$= 21.93). All participants were enrolled on a higher education degree at City University London, at the time of testing. In exchange for participating, individuals were awarded course credits, or a monetary reward of £8 per hour spent in the laboratory.

### 4.2.1.2 STIMULI AND PROCEDURE: READING TEST

Before carrying out the psychophysical experiments, all participants were administered 4 subtests from the Woodcock Reading Mastery Tests, 3rd edition (WRMT–III), which comprised the 'Word Attack', 'Word ID', 'Passage comprehension' and 'Reading Fluency' tests. The WRMT–III was standardised and validated in the United States on 3360 individuals aged 4-79 years.

#### 4.2.1.2.1 Word attack and Word ID

In the Word ID task, participants were required to read a series of 26 words presented to them on a computer screen. In the Word Attack participants were required to read a series of 22 nonsense words out loud, using the alphabetic and syllabication rules of the English language. The word attack and word ID both began with 4 practice items. Both subtests had a discontinue rule of four consecutive incorrect items, which did not have to be used on any of the participants. Raw scores consisted of the number of items read correctly and converted to standard scores according the WRMT conversion.

#### 4.2.1.2.2 Oral Reading Fluency

The Oral Reading Fluency subtest consisted of three passages which participants were required to read out loud. Reading scores were computed using the second and third passages only. Participants were instructed to read the passages in a natural voice, and were encouraged not to rush. The amount of errors the participant made was noted, as well as the time it took the participant to read the passage. To calculate the raw score for the Oral Reading Fluency, the following equation was used:

$$\left[\frac{(total\ word\ count - errors)}{time\ (sec.)}\right] \times 10 = passage\ raw\ score$$

The average raw score across the two passages were converted into a standard score, according the WRMT III conversion procedures.

#### 4.2.1.2.3 Passage comprehension

For the passage comprehension test, participants were required to silently read a series of passages which each contained a missing word. After reading each passage, the participant's task was to say out loud the word they believed to belong in the blank space. Only one-word responses were accepted. To calculate the raw score, all correct responses were added, and the total converted to a standard score according to the WRMT conversion.

### 4.2.1.3 STIMULI AND PROCEDURE: AV INTEGRATION AND TIMING TASKS

Tasks were carried out over two sessions, with a total duration of 140 minutes, inclusive of breaks. All participants carried out the McGurk Dual TOJ and Dual SJ tasks as well as the Stream bounce Dual TOJ and Dual SJ tasks, the stimuli and procedure for which are described in Chapter 2. The order in which these were performed was counterbalanced within groups.

### 4.2.1.4 ANALYSIS

Raw data from all tasks were handled in the same way as described in Chapter 3.

# 4.3 RESULTS: AUDIOVISUAL SPEECH INTEGRATION AND TIMING

One-tailed tests of comparison were used for analysing the difference between the groups for SD and Win measures for both implicit and explicit tasks, because it was predicted that dyslexic individuals would perform worse compared to controls, based on previous literature discussed in the introduction. Tests for all other measures (iPSS, ePSS, Max) were two-tailed as previous literature did not allow for any predictions regarding these. Partial correlations were carried out between aspects of reading ability (word identification, non-word reading, reading fluency and passage comprehension) and AV integration and temporal processing measures, controlling for dyslexia. These were all non-significant.

## 4.3.1 READING ABILITY ACROSS DYSLEXIA AND CONTROL GROUPS

Reading data from all participants were normally distributed within both the dyslexic and control groups and homogeneity of variance was met. Parametric statistics are shown in Table 4-1. Group comparisons of reading ability were carried out using parametric tests.

| | **Kolmorogov-Smirnov statistics** | | **Levene test statistics** |
|---|---|---|---|
| **Measure** | **Control group** | **Dyslexia Group** | |
| **Passage Comprehension** | *D*(23)=.11, *p*=.200 | *D*(19)=.19, *p*=.200 | *F*(1,40)=.00, *p*=.977 |
| **Reading Fluency** | *D*(23)=.12, *p*=.200 | *D*(19)=.16, *p*=.196 | *F*(1,40)=.25, *p*=.618 |
| **Word Attack** | *D*(23)=.17, *p*=.094 | *D*(19)=.13, *p*=.200 | *F*(1,40)=.12, *p*=.773 |
| **Word ID** | *D*(23)=.14, *p*=.200 | *D*(19)=.14, *p*=.200 | *F*(1,40)=.02, *p*=.881 |

*Table 4-1: Parametric assumption statistics for the reading ability variables.*

Reading ability differed significantly between dyslexic and control participants, with the exception of Passage comprehension [$t(40)=1.87$, $p=.069$]. Overall, the control group attained significantly higher scores in Word ID [$t(40)=4.03$, $p<.0005$], Word attack [$t(40)=4.69$, $p<.0005$] and reading fluency [$t(40)=4.07$, $p<.0005$] compared to dyslexic individuals. Means and standard deviations are shown in Table 4-2.

| | Typical readers (N=23) | | dyslexic readers (N=19) | |
|---|---|---|---|---|
| **Reading subtest** | **Mean** | **SD** | **Mean** | **SD** |
| **Passage Comprehension** | 101.96 | 9.88 | 96.47 | 8.94 |
| **Reading fluency** | 107.00 | 10.49 | 92.36 | 12.82 |
| **Word Attack** | 100.86 | 12.09 | 82.57 | 13.16 |
| **Word ID** | 106.00 | 8.20 | 95.36 | 8.85 |

*Table 4-2: Means and standard deviations of reading ability measures for the typical readers and the dyslexic groups.*

### 4.3.2 AUDIOVISUAL SPEECH SYNCHRONY JUDGEMENTS DUAL TASK (SJ)

Two participants (whose fits are illustrated in Figure 4.1) were excluded from the implicit temporal processing analysis due to having an inverse pattern of responses and/or unreliable fit.

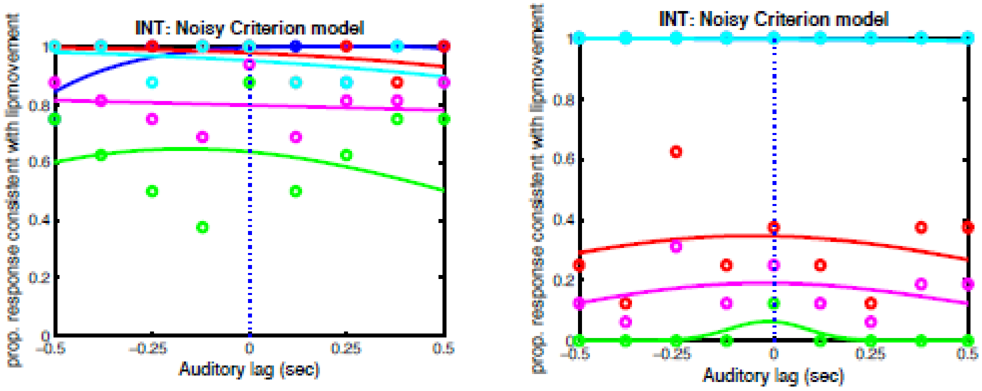


*Figure 4-1: Excluded McGurk integration data, fitted with noisy criterion fits. Left: inverted responses; Right: Flat function (SD>1).*

### 4.3.2.1 EXPLICIT SYNCHRONY JUDGEMENTS

The SD Data from synchrony judgements were normally distributed for the dyslexic [*W*(20)=.96, *p*=.551] but not for the control group [*W*(22)=.90, *p*=.036]. Homogeneity of variance was not met [*F*(1,40)=8.05, *p*=.007]. As a whole, the data were not normally distributed [*W(42)*=.93, *p*=.014]. Parametric statistics are shown in Table 4-3. Non-parametric tests were used to compare groups as well as correlate SJ SD measures with reading ability. All other explicit judgement measures met parametric assumptions and thus were analysed using parametric tests.

| **Measure** | | **Grouped stats** | | **Overall stats** |
|---|---|---|---|---|
| | | ***Shapiro-Wilk*** | ***Levene's Test*** | ***Shapiro-Wilk*** |
| **SJ SD** | **Control** | *W*(22)=.90, *p*=.036 | *F*(1,40)=8.05, *p*=.007 | *W(42)*=.93, *p*=.014 |
| | **Dyslexia** | *W*(20)=.96, *p*=.551 | | |
| **SJ Win** | **Control** | *W*(22)=.97, *p*=.607 | *F*(1,40)=0.37, *p*=.545 | *W*(42)=.98, *p*=.600 |
| | **Dyslexia** | *W*(20)=.97, *p=.775* | | |
| **SJ ePSS** | **Control** | *W*(22)=.97, *p*=.646 | *F*(1,40)=1.78, *p*=.190 | *W*(42)=.99, *p*=.950 |
| | **Dyslexia** | *W*(20)=.96, *p*=.606 | | |

*Table 4-3: Parametric assumptions for the explicit synchrony judgment measures: SJ SD, SJ Win and SJ ePSS*

**SJ SD:** The control group (Mdn = 0.13) on average showed significantly better performance in discriminating synchronous from asynchronous AV speech, with significantly smaller values of the SJ SD compared to the dyslexic group (Mdn=0.17) [*U*=128.00, *p*=.021*).* This difference is shown in Figure 4.2, on the next page.

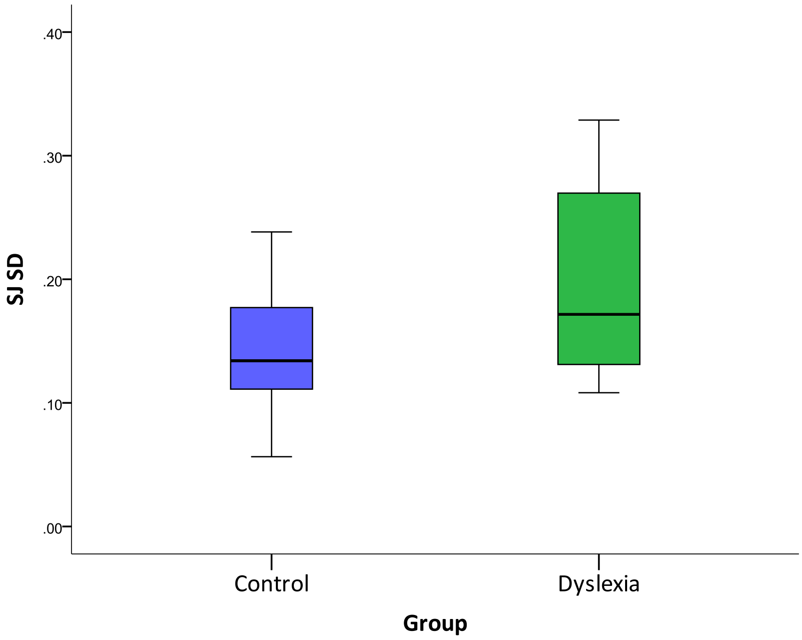


*Figure 4-2: Boxplot of SJ SD measures for the Control (blue) and Dyslexia (green) group. The dyslexia group on average showed significantly poorer performance in discrimination of AV synchrony (p=.021). Error bars represent 95% confidence intervals.*

**SJ PSS/SJ Win:** There were no significant differences between the two groups in terms of the SJ PSS or SJ Win. These measures were not significantly related to any of the reading ability measures. T-test statistics can be seen in Table 4-4, below.

| ***Measure*** | ***t test statistics*** |
|---|---|
| ***SJ PSS*** | *t*(40)= -.77, *p*=.446 |
| ***SJ Win*** | *t*(40)=-1.90, *p*=.065 |

*Table 4-4: (non-significant) t-test statistics for group differences in SJ PSS and SJ Win*

#### 4.3.2.2 IMPLICIT AUDIOVISUAL TIMING AND AUDIOVISUAL INTEGRATION (IN THE CONTEXT OF SJS)

All measures of implicit timing and AV integration met parametric assumptions, with the exception of the McG SD data from the control group. Parametric test statistics can be seen in Table 4-5, on the next page. All comparisons and correlations with the exception of the McG SD group comparisons were therefore run using parametric tests. Mann Whitney U test was used to compare group McG SD data.

| ***Measure*** | | ***Grouped stats*** | | ***Overall stats*** |
|---|---|---|---|---|
| | | ***Shapiro-Wilk*** | ***Levene's Test*** | ***Shapiro-Wilk*** |
| **McG SD** | **Control** | *W*(22)=.906, *p*=.047 | *F*(1,38)=1.179, *p*=.284 | *W(40)*=.946, *p*=.055 |
| | **Dyslexia** | *W*(19)=.954, *p*=.459 | | |
| **McG Win** | **Control** | *W*(22)=.932, *p*=.153 | *F*(1,38)=2.517, *p*=.121 | *W*(40)=.959, *p*=.157 |
| | **Dyslexia** | *W*(19)=.970, *p=.783* | | |
| **McG iPSS** | **Control** | *W*(22)=.934, *p*=.167 | *F*(1,38)=0.196, *p*=.661 | *W*(40)=.972, *p*=.412 |
| | **Dyslexia** | *W*(19)=.937, *p=.235* | | |
| **McG Max** | **Control** | *W*(22)=.981, *p*=.928 | *F*(1,39)=0.452, *p*=.505 | *W*(40)=.975, *p*=.512 |
| | **Dyslexia** | *W*(19)=.905, *p*=.059 | | |

*Table 4-5: Parametric assumptions tests for implicit AV timing measures (McG tMax, McG SD & McG Win) and AV integration (McG Max)*

**McGurk SD*:*** Measures of McGurk SD differed significantly between the dyslexic and Control groups [*U*=-.1.991*,* *p*=0.047], with dyslexic participants on average showing larger measures of the McGurk SD (Mdn =0.28) and thus poorer ability in implicit discrimination of synchrony compared to controls (Mdn= 0.23). This difference is illustrated in Figure 4.3, below.

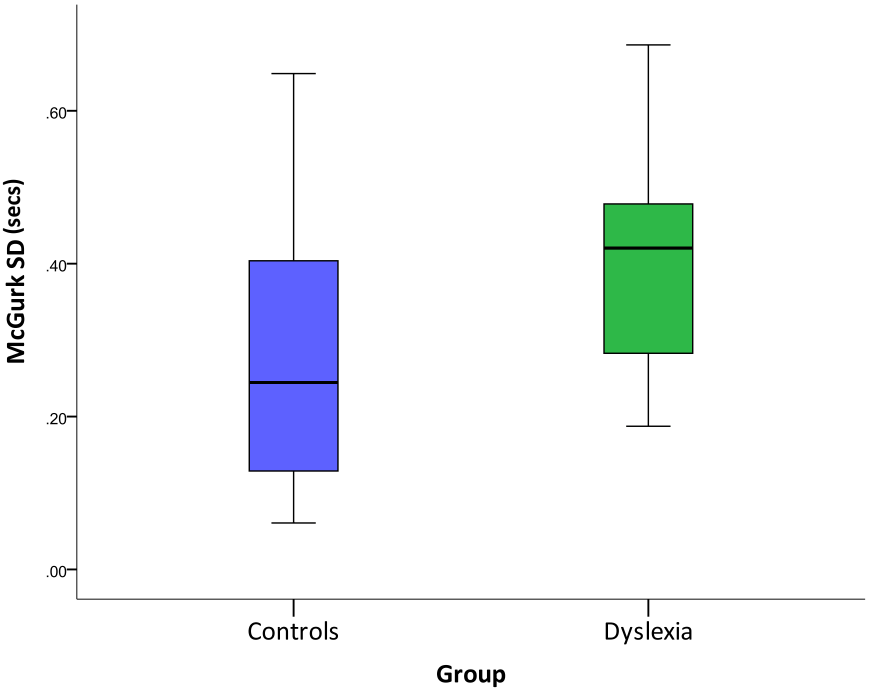


*Figure 4-3: Boxplot of McG SD measures for the Control (blue) and Dyslexia (green) group. The dyslexia group on average showed significantly poorer performance in discrimination of AV synchrony (p =0.047). Error bars represent 95% confidence intervals.*

**McGurk Win:** On average dyslexic participants (M=0.51, SD=0.23) showed a significantly wider window of AV speech integration, compared to controls (M=0.32, SD=0.19), [*t*(38)=2.85, *p*=.007]. This difference is shown in Figure 4.4, below.

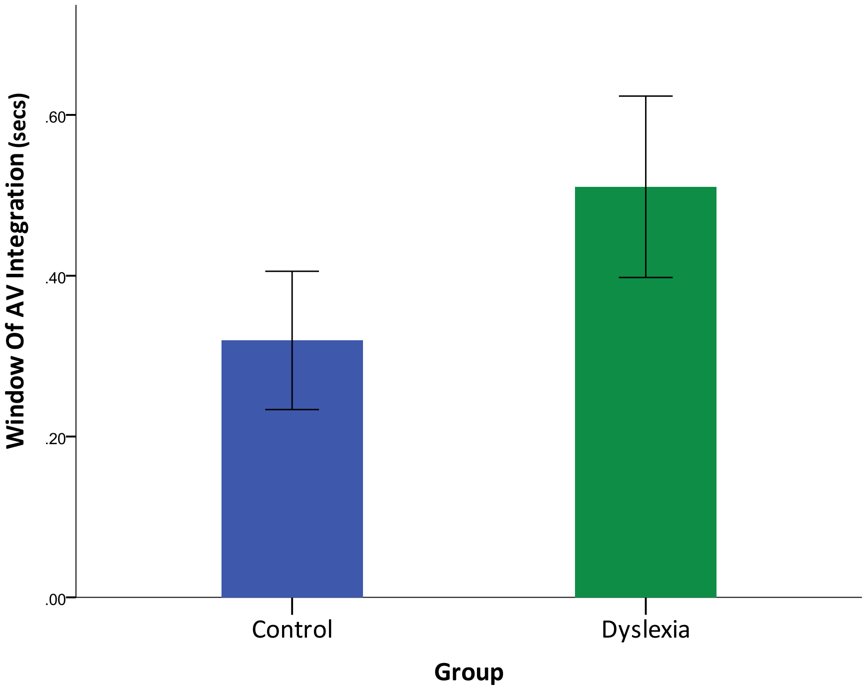


*Figure 4-4: Bar Chart of average measures of the window of AV Integration (McG Win) for the control (blue) and dyslexic (green) groups. The difference was significant (p=.007). Error bars represent 95% confidence intervals.*

**McG iPSS**: There was no significant group difference between the average AV asynchrony optimal for AV integration [t(39)=.401, *p*=691].

**McG Max:** Dyslexic individuals were on average significantly more susceptible (M=0.45, SD=0.23) to the McGurk effect compared to controls (M=0.63, SD=0.25) [*t*(39)=2.37, *p*=0.022], showing significantly larger proportion of illusory McGurk responses. This difference is shown in Figure 4.5, on the next page.

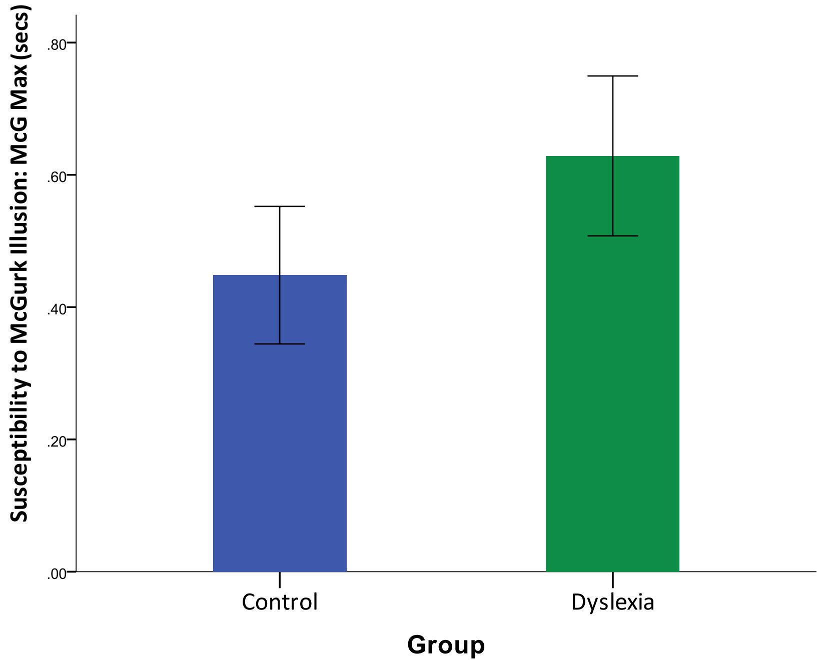


*Figure 4-5: Bar Chart of average measures of the maximum AV Integration (McG Max) for the control (blue) and dyslexic (green) groups. The difference was significant (p=.022). Error bars represent 95% confidence intervals.*

### 4.3.2.3 SUMMARY SJ DUAL TASK

Overall, dyslexic participants showed poorer explicit temporal discrimination and also on average more susceptible to the McGurk effect. They also showed on average significantly larger measures of the window of AV integration and significantly poorer implicit AV temporal discrimination.

### 4.3.3 AUDIOVISUAL SPEECH TEMPORAL ORDER JUDGEMENTS DUAL TASK (TOJ)

Twenty five percent of participants from each group were excluded from the TOJ analysis due to unreliable fits as a result of chance performance at every level of AV asynchrony or inverted psychometric functions (see Figure 4.6 for example).

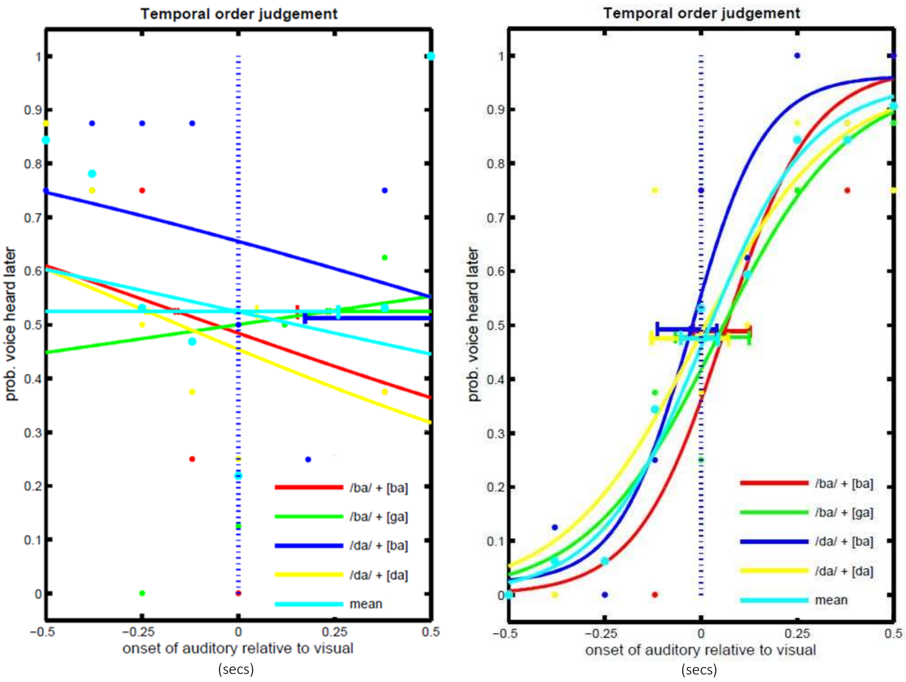


*Figure 4-6: Example of unreliable TOJ data (left) and example of reliable TOJ data (right)*

#### 4.3.3.1 EXPLICIT TEMPORAL ORDER JUDGEMENTS

The TOJ SD data from temporal order judgements were normally distributed for both groups and homogeneity of variance was met. Parametric statistics can be seen in Table 4-6, on the next page. As a whole, the data were normally distributed. Parametric tests were used to compare groups as well as to correlate TOJ SD measures with reading ability. TOJ ePSS measures met parametric assumptions and thus were analysed using parametric tests.

| Measure | | Grouped stats | | Overall stats |
|---|---|---|---|---|
| | | ***Shapiro-Wilk*** | ***Levene's Test*** | ***Shapiro-Wilk*** |
| **TOJ SD** | **Control** | *W*(18)=.92, *p*=.139 | *F*(1,31)=1.51, *p*=.228 | *W(33)*=.96, *p*=.342 |
| | **Dyslexia** | *W*(15)=.99, *p*=.993 | | |
| **TOJ ePSS** | **Control** | *W*(18)=97, *p*=.741 | *F*(1,31)=0.02, *p*=.890 | *W*(33)=.96, *p*=.242 |
| | **Dyslexia** | *W*(15)=.90, *p*=.108 | | |

*Table 4-6: Parametric assumptions tests for explicit AV temporal order judgement measures (TOJ)*

**TOJSD:** The control group (M=0.50, SD=0.20) showed on average better performance at discriminating AV temporal order, with significantly smaller values of the TOJ SD compared to the dyslexic group (M=0.65, SD=0.27) [*t(31)*=-1.89, *p*=0.034, one-tailed). This difference is illustrated below, in Figure 4.7.

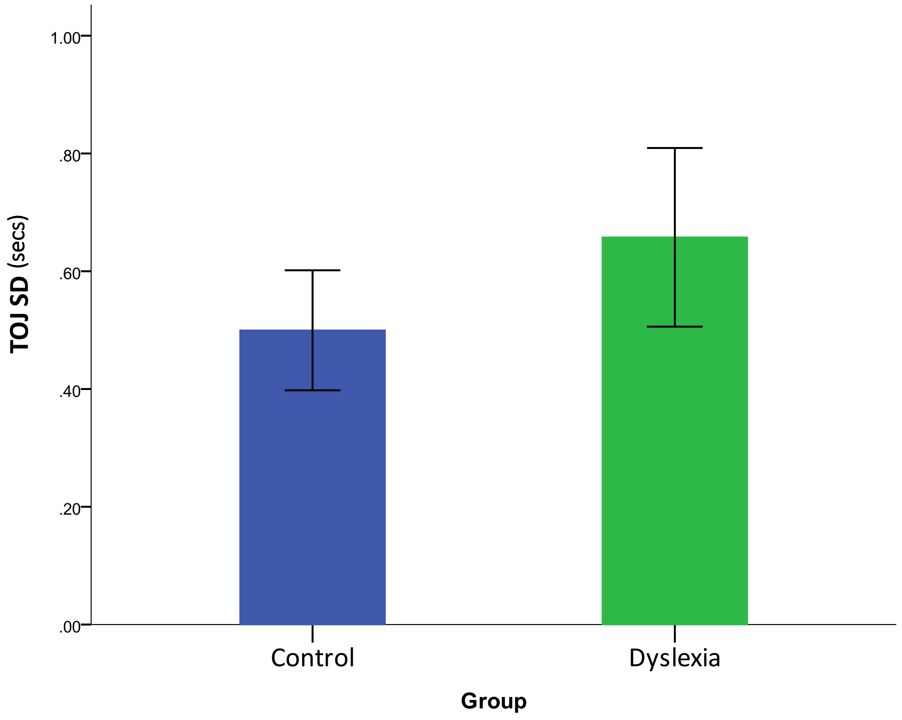


*Figure 4-7: Bar chart of average measures of TOJ SD for the Control (Blue) and dyslexia (green) groups. The difference was significant (p=.034). Error bars represent 95% confidence intervals.*

**TOJ ePSS:** TOJ ePSS measures did not differ significantly between the two groups [*t*(31)=-0.46, *p*=.652].

### 4.3.3.2 IMPLICIT AUDIOVISUAL TIMING AND AUDIOVISUAL INTEGRATION (IN THE CONTEXT OF TOJS)

All measures of implicit timing and AV integration met parametric assumptions, with the exception of measures of the Window of AV Speech Integration from the control group, which were not normally distributed. Statistics can be seen in Table 4-7, below. Parametric tests were used to compare measures of McG SD, iPSS and susceptibility to the McGurk illusion between the groups, and a non-parametric comparison test was used for the Window of AV Integration.

| ***Measure*** | ***Grouped stats*** | | | ***Overall stats*** |
|---|---|---|---|---|
| | | ***Shapiro-Wilk*** | ***Levene's Test*** | ***Shapiro-Wilk*** |
| **McG iPSS** | **Control** | *W*(24)=.965, *p*=.550 | *F*(1,42)=0.081, *p*=.284 | *W(43)*=.979, *p*=.619 |
| | **Dyslexia** | *W*(10)=.984, *p*=.972 | | |
| **McG Max** | **Control** | *W*(24)=.963, *p*=.511 | *F*(1,42)=0.541, *p*=.466 | *W*(43)=.967, *p*=.247 |
| | **Dyslexia** | *W*(20)=.953, *p=.417* | | |
| **McG Win** | **Control** | *W*(24)=.743, *p<.001* | *F*(1,42)=0.479, *p*=.493 | *W*(43)=.883, *p*<.001 |
| | **Dyslexia** | *W*(20)=.929, *p=.145* | | |
| **McG SD** | **Control** | *W*(24)=.942, *p*=.183 | *F*(1,42)=1.031, *p*=.316 | *W*(43)=.922, *p*=.006 |
| | **Dyslexia** | *W*(20)=.895, *p*=.050 | | |

*Table 4-7: Parametric assumptions tests for implicit AV timing measures (tMax, McG SD and McG Win) and AV integration (McG Max)*

**McGurk SD:** Dyslexic individuals (M=0.32, SD=0.13) on average had larger measures of the McGurk SD compared to controls (M=0.26, SD=0.11) and thus showed poorer ability to implicitly discriminate between synchronous and asynchronous AV speech information when integrating AV information. This difference was borderline significant [*t*(42)=-1.65, *p*=.054, one-tailed] and can be seen in Figure 4.8, on the next page.

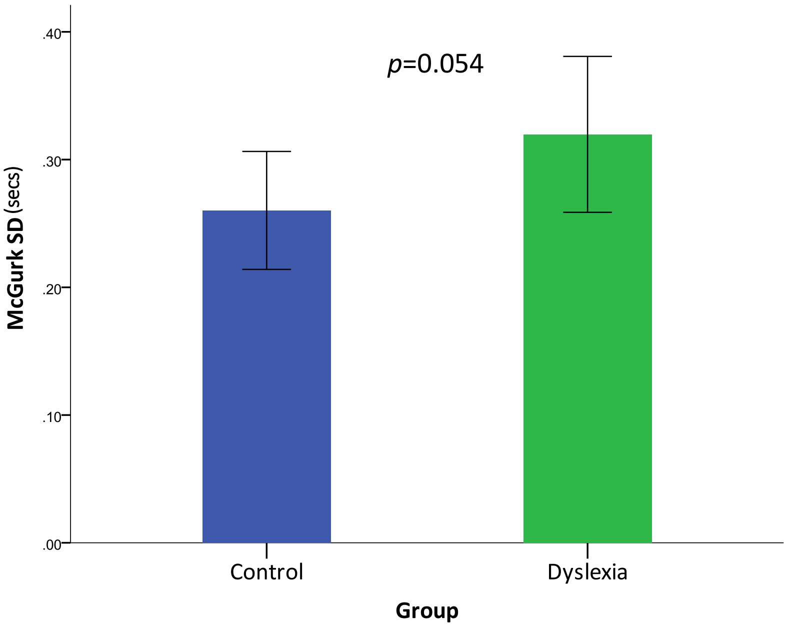


*Figure 4-8: Bar chart of average measures of McGurk SD for the control (blue) and dyslexia (green) group. The difference was borderline significant (p=0.054). Error bars represent 95% confidence intervals.*

**McG Win:** On average dyslexic participants (Mdn=0.28) had significantly wider windows of AV speech integration, compared to controls (Mdn=0.23) [*U=138.0, p*=0.008, one-tailed] (see Figure 4.9 below), meaning that on average they integrate AV speech information at significantly larger AV asynchronies.

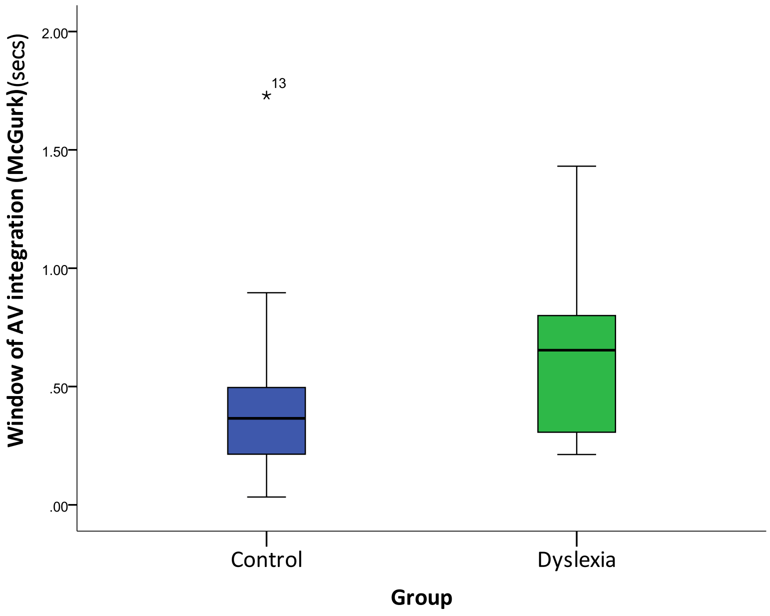


*Figure 4-9: Box plot for average measures of McG Win for Control (blue) and dyslexia (green) group. The difference was significant (p=0.008). Error bars represent 95% confidence intervals.*

**McGurk iPSS:** There was no significant group difference between the average AV asynchrony optimal for AV integration [t(43)=-0.354, *p*=.725, two tailed].

**McG Max:** Dyslexic participants (M=0.65, SD=0.18) were on average significantly more susceptible to the McGurk illusion compared to controls (M=0.51, SD=0.24) [*t*(42)=2.13, *p*=0.039]. This difference is shown in Figure 4.10, below.

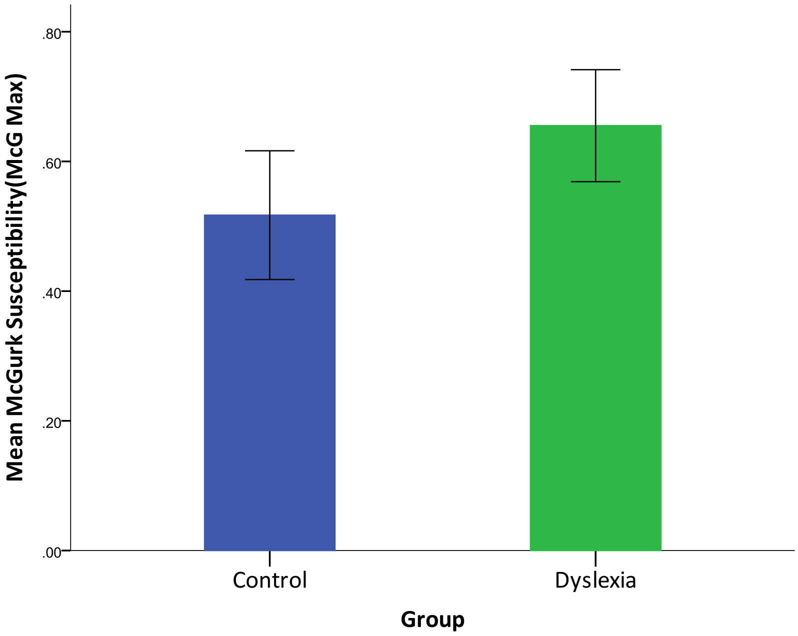


*Figure 4-10: Bar chart of average measures of McG Max for Control (blue) and dyslexia (green) group. The difference was significant (p=0.039). Error bars represent 95% confidence intervals.*

#### 4.3.3.3 SUMMARY AUDIOVISUAL TOJ AND AUDIOVISUAL SPEECH INTEGRATION DUAL TASK

On average dyslexic participants had poorer ability in explicit AV temporal order discrimination, shown by larger measures of the TOJ SD. Following this trend, there was also a borderline significant difference between the two groups in terms of the implicit ability to discriminate synchronous from asynchronous AV speech information. The dyslexic group on average also showed significantly larger windows of AV integration. Dyslexic participants were significantly more susceptible to the McGurk illusion.

# 4.4 RESULTS: AUDIOVISUAL NON-SPEECH: STREAM BOUNCE

## 4.4.1 READING ABILITY ACROSS DYSLEXIA AND CONTROL GROUPS

Reading data from all participants were normally distributed within both the dyslexic and control groups and homogeneity of variance was met. Parametric statistics can be seen in Table 4-8, below. Group comparisons of reading ability were carried out using parametric tests.

| | Kolmorogov-Smirnov statistics | | Levene test statistics |
|---|---|---|---|
| **Measure** | **Control group** | **Dyslexia Group** | |
| **Passage Comprehension** | *W*(22)=.95, *p*=.365 | *W*(18)=.96, *p*=.562 | *F*(1,38)=.00, *p*=.987 |
| **Reading Fluency** | *W*(22)=.96, *p*=.447 | *W*(18)=.92, *p*=.105 | *F*(1,38)=.10, *p*=.758 |
| **Word Attack** | *W*(22)=.92, *p*=.086 | *W*(18)=.98, *p*=.968 | *F*(1,38)=.00, *p*=.975 |
| **Word ID** | *W*(22)=.94, *p*=.198 | *W*(18)=.94, *p*=.321 | *F*(1,38)=.01, *p*=.935 |

*Table 4-8: Parametric assumption statistics for reading ability data for control and dyslexic groups.*

Reading ability differed significantly between dyslexic and control participants, with the exception of Passage comprehension [*t*(38)=1.50*, p*=.142]. Overall, the control group attained significantly higher scores in Word ID [*t*(38)=4.32, *p*<.0005], Word attack [*t*(38)=5.49, *p*<.0005] and reading fluency [*t*(38)=4.09, *p*<.0005] compared to dyslexic individuals. Means and standard deviations for reading ability by group are shown in Table 4-9, below.

| | Typical readers (N=22) | | dyslexic readers (N=18) | |
|---|---|---|---|---|
| **Reading subtest** | **Mean** | **SD** | **Mean** | **SD** |
| **Passage Comprehension** | 101.50 | 9.86 | 97.00 | 8.99 |
| **Reading fluency** | 107.00 | 10.74 | 91.72 | 12.87 |
| **Word Attack** | 101.45 | 12.04 | 82.57 | 11.38 |
| **Word ID** | 106.36 | 8.20 | 94.78 | 8.71 |

*Table 4-9: Means and standard deviations for reading ability measures, for the control and dyslexia groups.*

## 4.4.2 Audiovisual Non-Speech Synchrony Judgements Dual task (SJ)

### 4.4.2.1 Explicit Synchrony Judgements

Data were normally distributed for both groups and homogeneity of variance was met. Overall, data were normally distributed for SJ WIN and SJ PSS, but not for SJ SD. All analyses were carried out using parametric tests. Statistics are shown in Table 4-10, below.

| **Measure** | | **Grouped stats** | | **Overall stats** |
|---|---|---|---|---|
| | | ***Shapiro-Wilk*** | ***Levene's Test*** | ***Shapiro-Wilk*** |
| **SJ SD** | **Control** | *W*(22)=.93, *p*=.142 | *F*(1,38)=2.80, *p*=.102 | *W(40)*=.94, *p*=.025 |
| | **Dyslexia** | *W*(20)=.92, *p*=.123 | | |
| **SJ Win** | **Control** | *W*(22)=.98, *p*=.896 | *F*(1,38)=0.79, *p*=.381 | *W*(40)=.10, *p*=.964 |
| | **Dyslexia** | *W*(20)=.96, *p*=.654 | | |
| **SJ PSS** | **Control** | *W*(22)=.97, *p*=.598 | *F*(1,40)=0.08, *p*=.781 | *W*(40)=.97, *p*=.291 |
| | **Dyslexia** | *W*(20)=.95, *p*=.361 | | |

*Table 4-10: Parametric assumptions test statistics for explicit AV synchrony measures (SJ)*

**SJ SD:** No significant difference was found between the dyslexic group (M=0.15, SD=0.09) and controls (M=0.12, SD=0.06) in the ability to discriminate between synchronous and asynchronous AV non-speech information [*t*(38)=-1.21, *p*=.117, one-tailed].

**SJ ePSS**: There were no significant differences between the dyslexic (M=0.02, SD=0.05) and control (M=-0.01, SD=0.05) groups in terms of the SJ PSS [*t*(38)=-1.78, *p*=0.083, two tailed).

**SJ Win:** There were no significant differences between the dyslexic (M=0.41, SD=0.09) and control (M=0.34, SD=0.16) groups in terms of the size of the Window of AV synchrony [*t*(38)=-1.46, *p*=0.076, one-tailed).

#### 4.4.2.2 Stream-Bounce Integration: Implicit AV timing and AV integration (in the context of SJs)

Overall, 36.4% of controls and 44.4% of dyslexic participants were excluded from the analysis due to not experiencing the Stream-Bounce illusion, or having flat functions (see Figure 4.11 below for one example). There was no significant association between group and exclusion rate [$x^2$(1)=.269, *p*=.604].

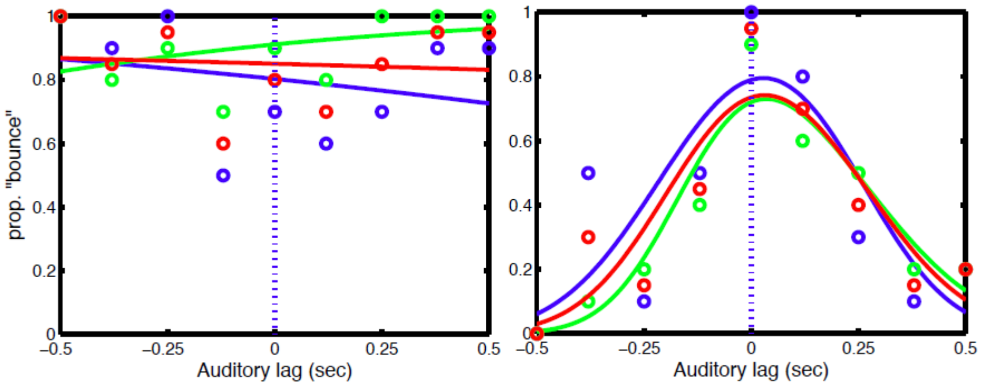


*Figure 4-11: Example of unreliable Stream-Bounce data which seems to have an inverse function (right) and typical Stream-Bounce data (left).*

All measures of implicit timing and AV integration met parametric assumptions. Parametric statistics are shown below in Table 4-11. All analyses were carried out using parametric tests.

| ***Measure*** | | ***Grouped stats*** | | ***Overall stats*** |
|---|---|---|---|---|
| | | ***Shapiro-Wilk*** | ***Levene's Test*** | ***Shapiro-Wilk*** |
| **Stream-Bounce SD** | **Control** | *W*(14)=.92, *p*=.200 | *F*(1,22)=0.36, *p*=.553 | *W*(24)=.95, *p*=.267 |
| | **Dyslexia** | *W*(10)=.94, *p*=.530 | | |
| **Stream-Bounce Win** | **Control** | *W*(14)=.98, *p*=.959 | *F*(1,22)=0.71, *p*=.410 | *W*(24)=.98, *p*=.923 |
| | **Dyslexia** | *W*(10)=.97, *p=.891* | | |
| **Stream-Bounce iPSS** | **Control** | *W*(14)=.90, *p*=.107 | *F*(1,22)=0.76, *p*=.392 | *W*(24)=.94, *p*=.175 |
| | **Dyslexia** | *W*(10)=.95, *p*=.699 | | |
| **Bounce Max** | **Control** | *W*(14)=.96, *p*=.669 | *F*(1,22)=0.08, *p*=.774 | *W*(24)=.97, *p*=.566 |
| | **Dyslexia** | *W*(10)=.97, *p*=.587 | | |

*Table 4-11: Parametric assumptions tests for implicit AV timing measures and AV integration*

**Stream Bounce SD:** Measures of Stream-Bounce SD differed significantly between the dyslexic and Control groups [*t(22)*=-1.77, *p*=0.046, one-tailed], with dyslexic participants on average showing larger measures of the SD (M=0.42, SD=0.17) and thus poorer implicit AV temporal discriminability compared to controls (M=0.29, SD=0.19). This difference can be seen below in Figure 4.12.

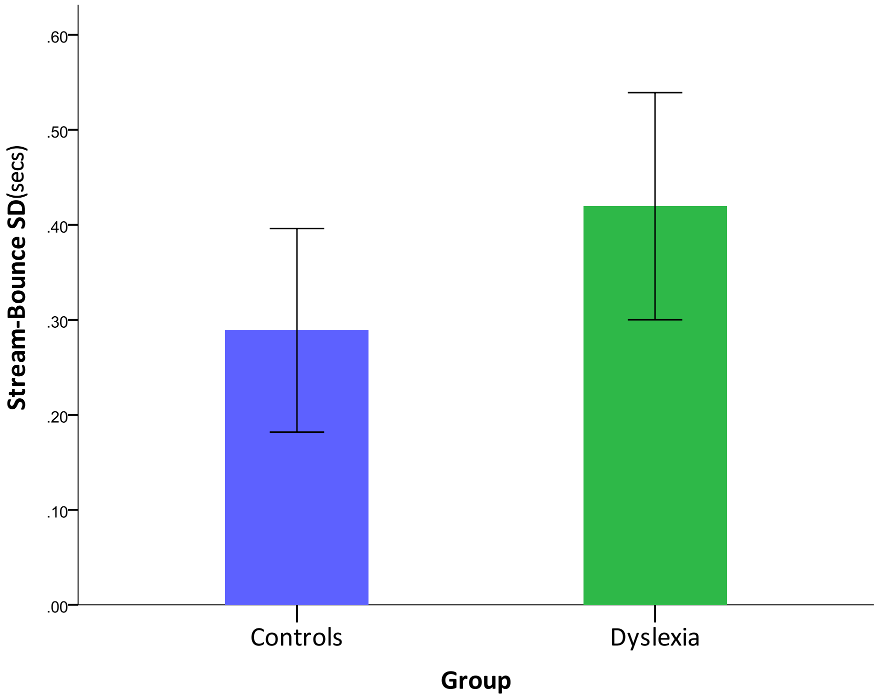


*Figure 4-12: Bar chart of average measures of Stream-Bounce SD for Control (blue) and dyslexia (green) group. The difference was significant (p=0.046). Error bars represent 95% confidence intervals.*

**Stream Bounce Win:** The dyslexic (M=0.55, SD=0.27) and control (M=0.50, SD=0.22) did not differ significantly in terms of the Window of AV non-speech integration, [*t*(22)=-0.48, *p*=.637].

**Stream-Bounce iPSS:** There was no significant group difference between the average AV asynchrony optimal for AV integration [t(22)=0.76, *p*=.456, two tailed].

**Bounce Max:** There was no significant difference between the groups in terms of susceptibility to the Stream-Bounce illusion (Bounce Max) [*t*(22)=1.49, *p*=0.150].

#### 4.4.2.3 SUMMARY NON-SPEECH AV SJ AND AV INTEGRATION DUAL TASK

Dyslexic participants showed on average showed poorer ability to implicitly discriminate between synchronous and asynchronous AV stimuli during AV integration, shown by significantly larger measures of the Stream-Bounce SD. Overall, there were on average no differences between the two groups in terms of explicit non-speech synchrony processing. The windows of AV integration and AV synchrony were equivalent across the two groups. No differences were found in terms of susceptibility to the Stream-Bounce illusion between the two groups.

### 4.4.3 AUDIOVISUAL NON-SPEECH TEMPORAL ORDER JUDGEMENTS DUAL TASK (TOJ)

#### 4.4.3.1 EXPLICIT TEMPORAL ORDER JUDGEMENTS

One control participant (4.5%) and 3 dyslexic participants (16.7%) were excluded from the TOJ analysis due to poor fits leading to extreme measures (PSS>0.4). There was no significant association between group and exclusion rate [$x^2$(1)=1.61, *p*=.204].

Data were normally distributed for both groups and homogeneity of variance was met. Overall, data were normally distributed for all measures. Parametric statistics are shown below, in Table 4-12. All analyses were carried out using parametric tests.

| **Measure** | | **Grouped stats** | | **Overall stats** |
|---|---|---|---|---|
| | | ***Shapiro-Wilk*** | ***Levene's Test*** | ***Shapiro-Wilk*** |
| **TOJ SD** | **Control** | *W*(21)=.94, *p*=.252 | *F*(1,34)=1.73, *p*=.198 | *W*(*36*)=.97, *p*=.347 |
| | **Dyslexia** | *W*(25)=.94, *p*=.343 | | |
| **TOJ ePSS** | **Control** | *W*(21)=.97, *p*=.681 | *F*(1,34)=0.07, *p*=.795 | *W*(36)=.98, *p*=.868 |
| | **Dyslexia** | *W*(15)=.91, *p*=.145 | | |

*Table 4-12: Parametric assumption statistics for TOJ SD and TOJ ePSS*

**TOJ SD:** No significant difference was found between the groups (Dyslexia M=0.25, SD=0.09; Control M=0.22, SD=0.12) in the ability to discriminate AV non-speech temporal order [*t*(34)=-0.94, *p*=.354, one-tailed].

**TOJ ePSS:** There were no significant differences between the two groups (Dyslexia M=0.06, SD =0.08; Control M=0.02, SD=0.09) in terms of the TOJ ePSS [*t*(34)=-1.44, *p*=.158, two tailed).

#### 4.4.3.2 STREAM-BOUNCE INTEGRATION: IMPLICIT AV TIMING AND AV INTEGRATION (IN THE CONTEXT OF TOJS)

Overall, 36.4% of controls and 33.3% of dyslexic participants were excluded from the analysis due to not experiencing the Stream-Bounce illusion, or having flat functions. There was no significant association between group and exclusion rate [$x^2$(1)=0.04*, p*=.842].

All measures of implicit timing and AV integration met parametric assumptions. Parametric tests statistics are shown below in Table 4-15. All analyses were carried out using parametric tests.

| ***Measure*** | | ***Grouped stats*** | | ***Overall stats*** |
|---|---|---|---|---|
| | | ***Shapiro-Wilk*** | ***Levene's Test*** | ***Shapiro-Wilk*** |
| **Stream Bounce SD** | **Control** | *W*(14)=.92, *p*=.171 | *F*(1,24)=0.32, *p*=.577 | *W*(26)=.96, *p*=.292 |
| | **Dyslexia** | *W*(10)=.94, *p*=.786 | | |
| **Stream-Bounce Win** | **Control** | *W*(14)=.98, *p*=.786 | *F*(1,24)=2.45, *p*=.131 | *W*(26)=.95, *p*=.867 |
| | **Dyslexia** | *W*(10)=.97, *p*=.772 | | |
| **Stream-Bounce iPSS** | **Control** | *W*(14)=.90, *p*=.152 | *F*(1,24)=0.31, *p*=.580 | *W*(26)=.98, *p*=.749 |
| | **Dyslexia** | *W*(10)=.95, *p*=.618 | | |
| **Bounce Max** | **Control** | *W*(14)=.96, *p*=.670 | *F*(1,24)=2.13, *p*=.158 | *W*(26)=.96, *p*=.485 |
| | **Dyslexia** | *W*(10)=.97, *p*=.539 | | |

*Table 4-13: Parametric assumptions statistics for AV integration and implicit timing measures.*

**Stream-Bounce SD:** Stream-Bounce SD measures differed significantly between the dyslexic and Control groups [*t*(24)=-2.99*, p*=.003, one-tailed], with dyslexic participants on average showing larger Stream-Bounce SD measures (M=0.42, SD=0.13) and thus poorer ability in implicit discrimination of AV synchrony from asynchrony compared to controls (M=0.27, SD=0.14). this difference is illustrated in Figure 4.13, on the next page.

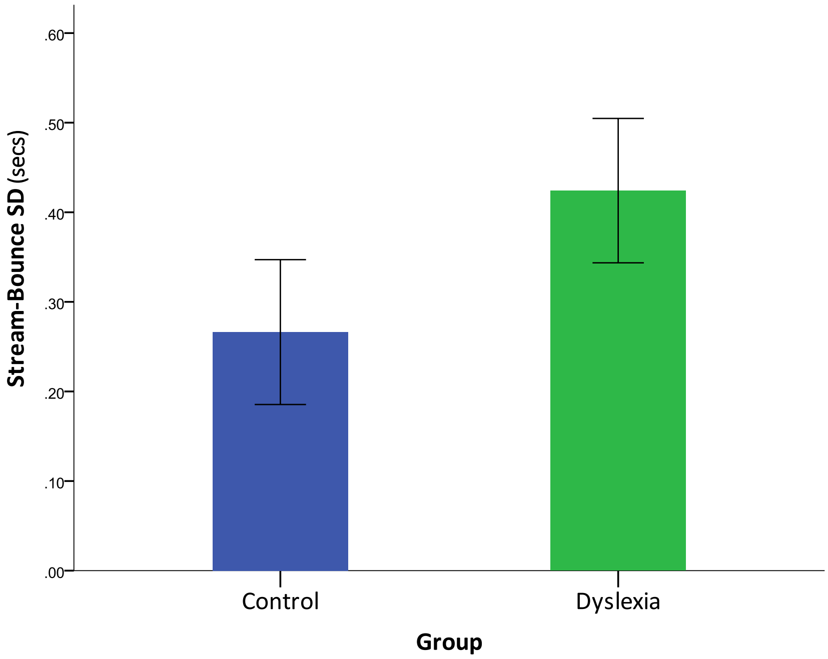


*Figure 4-13: Bar chart of average measures of Stream-Bounce SD for the control (blue) and dyslexic Group. The difference was significant (p=.003). Error bars represent 95% confidence intervals.*

**Stream-Bounce Win: T**he two groups (Dyslexia M=0.49, SD =0.28; Control M=0.52, SD=0.19) did not differ significantly in terms of the Window of AV non-speech integration [*t*(24)=0.37, *p*=.356, one-tailed].

**Bounce iPSS:** There was no significant group difference between the average AV asynchrony optimal for AV integration (iPSS) [t(22)=-0.18, *p*=.857, two tailed].

**Bounce Max:** There was a significant difference between the groups in terms of susceptibility to the Stream-Bounce illusion [*t*(24)=3.10, *p*=0.005, two-tailed], with controls showing on average higher susceptibility (M=0.70, SD=0.17) relative to dyslexic participants (M=0.46, SD=0.24). This difference is illustrated in Figure 4.14, on the next page.

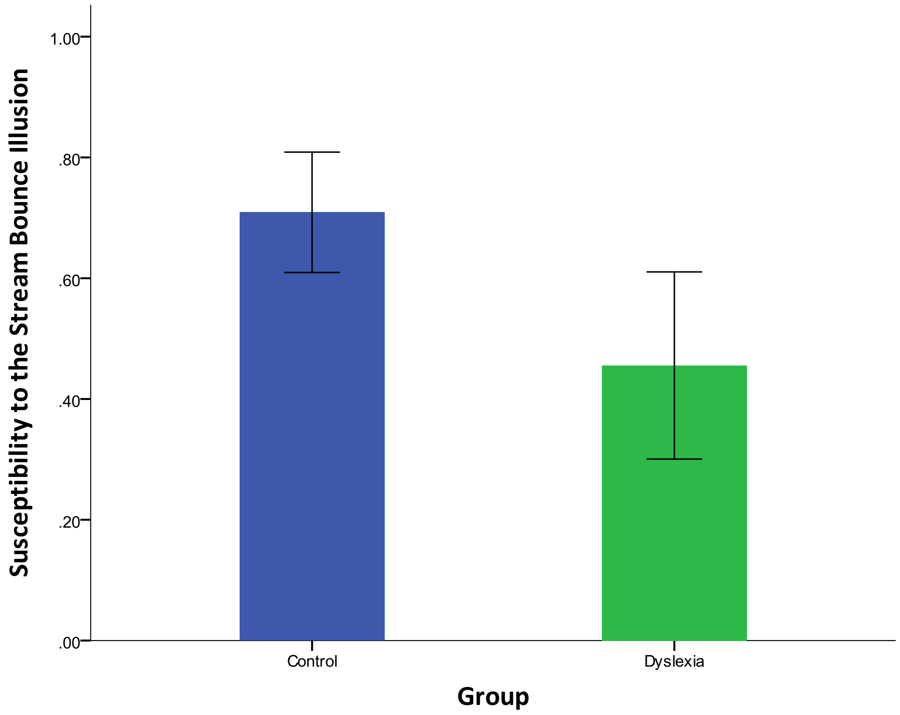


*Figure 4-14: Bar charts of average measures of Bounce Max for the control (blue) and dyslexia (green) groups. Error bars represent 95% confidence intervals.*

### 4.4.3.3 SUMMARY NON-SPEECH TOJ AND AV INTEGRATION DUAL TASK

There were no differences between the two groups in terms of explicit judgements of AV temporal order. Dyslexic participants however on average showed poorer ability to implicitly discriminate between synchronous and asynchronous AV stimuli during AV integration, shown by significantly larger measures of the implicit JND. The window of AV integration was equivalent across the two groups. Dyslexic individuals also showed lower rates of AV integration, shown by lower susceptibility to the Stream bounce illusion.

## 4.5 ADDITIONAL ANALYSES USED FOR INTERPRETATION

The relationship between the ability to implicitly discriminate AV synchrony from asynchrony and individual rates of AV integration was measured across all the tasks. In the two McGurk tasks, there was no significant relationship between the McG SD and McG Max measures. Within the Stream-Bounce task, poorer discrimination ability was related to less AV integration (see Figure 4.15 and 4-16 for relationships), both in the TOJ dual-task [*r*(26)=-.649, *p*<.0005] and the SJ Dual task [*r*(26)=-.620, *p*=.001].

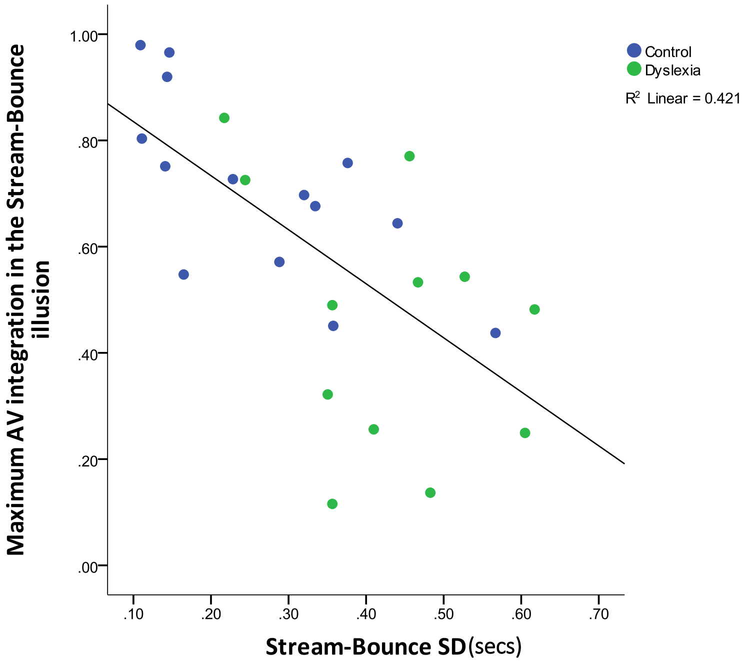


*Figure 4-15: Scatter plot of significant negative correlation between Stream-Bounce SD and Stream-Bounce Max in the TOJ Dual task.*

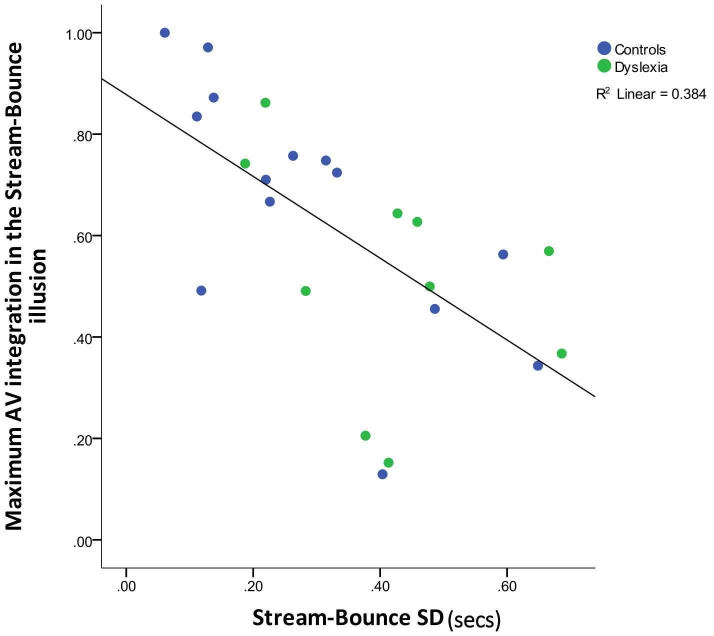


*Figure 4-16: Scatter plot of significant negative correlation between Stream-Bounce SD and Stream-Bounce Max in the TOJ Dual task*

## 4.6 RESULTS SUMMARY

Overall, dyslexic participants showed poorer explicit temporal discrimination of speech stimuli in both TOJs and SJs. Across both speech dual-tasks, dyslexic individuals were also on average more susceptible to the McGurk effect. They also showed on average significantly larger measures of the window of AV integration and significantly poorer implicit AV temporal discrimination.

In the non-speech TOJ and SJ task, dyslexic participants showed no impairments in explicit discrimination of AV synchrony and temporal order. Performance was poorer in the dyslexic group for implicit discrimination of AV synchrony. AV integration levels were equivalent across groups in the SJ non-speech dual-task and lower for the dyslexic group in the TOJ non-speech dual-task. Reading ability was not related to AV integration and timing, over and above dyslexia. A summary of the results is also shown in Table 4-14 on the next page.

|  |  | **Explicit temporal processing** |  | **Implicit temporal Processing** |  | **AV Integration** |
|---|---|---|---|---|---|---|
|  |  | SD | Win | SD | Win | Max |
| **Speech** | SJ | D>C | D=C | D>C | D>C | D>C |
|  | TOJ | D>C | ---- | D>C<br>*p=.054* | D>C | D>C |
| **Non-speech** | SJ | D=C | D=C | D>C | D=C | D=C |
|  | TOJ | D=C | ---- | D>C | D=C | D<C |

*Table 4-14: Summary of Chapter 4 results. D: Dyslexia C: Controls*

# 4.7 DISCUSSION

## 4.7.1 PERFORMANCE IN EXPLICIT AUDIOVISUAL TEMPORAL PROCESSING IS POORER IN DYSLEXIC READERS FOR SPEECH BUT NOT FOR NON-SPEECH STIMULI

Dyslexic individuals on average showed significantly poorer ability in discriminating synchronous from asynchronous AV speech and worse performance in discriminating the temporal order of AV speech. This difference was not found for explicit temporal processing of non-speech stimuli. Within speech processing literature, dyslexic individuals have only so far been reported to show poorer ability in auditory speech temporal order judgements (De Martino et al., 1999; Rey et al., 2002). The current investigation is the first to indicate that explicit temporal processing of speech is also impaired *across* modalities in dyslexia. In SJ profiles, a wider window of synchrony can be interpreted as a bias towards responding that the AV information was presented synchronously, or having more relaxed decision making criteria. The window of AV speech synchrony was equivalent across the two groups suggesting dyslexic group did not have more relaxed decision making criteria regarding what is synchronous and what is not, or a general bias towards responding that the AV stimuli were synchronous.

Poorer explicit temporal processing ability in dyslexia was not found in Experiments 9 and 10, in which explicit temporal judgements were made on non-speech stimuli; dyslexic performance in explicit discrimination of AV non-speech temporal order and in discriminating synchronous from asynchronous AV non-speech stimuli was statistically equivalent to that of controls. This result concurs with findings reported by Laasonen et al. (2002), who reported that although dyslexic individuals exhibited poorer temporal sensitivity in audiotactile and visuotactile temporal judgements, their performance was normal in AV TOJs and SJs using non-speech AV stimuli consisting of sound bursts and flashes. Altogether, these results suggest that explicit temporal processing impairments in dyslexia may be speech specific.

It is unlikely that the results can be accounted for by the notion that difficulties in temporal processing of speech in dyslexia stem from impaired processing of rapidly occurring features of auditory speech, such as short consonant sounds which occur in rapid succession (Raschle, Stering, Meissner, & Gaab, 2013; Tallal, 1980a). This is because the speech stimuli used here do not meet the criteria for what constitutes as rapidly occurring speech sounds (Raschle et al., 2013; Tallal, 1980a); each auditory stimuli here comprised of only one consonant followed by one vowel. Poorer explicit discrimination of AV temporal order and synchrony in AV speech cannot be accounted by an exclusive impairment in phonological processing either. To recap, phonological theories of dyslexia argue that reading impairment in the disorder is caused exclusively by left hemisphere dysfunction which leads to deficient phonological representations, in turn affecting grapheme-phoneme mapping and other phonological skills (Stanovich, 1988). Such impairment might contribute towards poorer ability to identify which phoneme was heard in the phoneme ID task, but not towards poorer discrimination of the relative onset timing of auditory and visual speech.

One example of how poor explicit temporal processing of AV speech might contribute to reading impairments in dyslexia is by affecting the ability to read words in a bottom-up fashion. Reading novel, or non-words more likely requires that individual graphemes of words are translated one by one into their auditory representations and then combined. High sensitivity to AV temporal order might be useful for explicitly monitoring the onset of speech sound representations relative to the written letters which activate them during decoding of individual graphemes within novel words. This in turn might be beneficial for mapping speech representations order across the two modalities to ensure that phonemes are combined in the auditory modality into a single word according to the order in which they are represented visually on the page.

The ability to discriminate AV synchrony from asynchrony as well as to discriminate AV temporal order might also affect reading ability such as reading fluency. Familiar word identification is likely to recruit top down processes, that is, recognising entire words which the reader has encountered before and retrieving the auditory representation of that word rather than decoding its individual graphemes. Passages of text contain both familiar and novel words, meaning that reading fluency will likely recruit both top down and bottom up reading processes. When reading entire passages, the ability to synchronise and discriminate the order of AV speech might contribute to inhibiting irrelevant sensory information, for example other auditory representations activated by words surrounding the one that is being read or attended to (Hairston et al., 2005), in turn maintaining reading fluency, by minimising hesitations and maximising reading speed.

Audiovisual temporal processing is unlikely to explain all reading impairments in dyslexia. For example, the above discussion of how discrimination of AV speech temporal order might contribute to reading can to some extent explain mistakes characterised by letter reversal, as well as slower reading. It cannot however straightforwardly explain mistakes characterised by poor formation of grapheme-phoneme correspondences, such as retrieval of the incorrect phoneme for a given grapheme, for example pronouncing /ĕ/ (for m**e**) instead of /ī/ in the word ‘fly’.

In Chapter 3, better performance in explicit AV speech synchrony and temporal order tasks was related to greater density of grey matter in the right superior temporal cortex. The temporal cortex is one of the areas associated with functional (Blau et al., 2010, 2009) and structural (Eckert et al., 2005) abnormalities in dyslexia. Widespread functional or structural abnormalities are likely to lead to impairments in a range of processes supported by these brain areas, including temporal processing. One possibility is that abnormal development of brain structure in dyslexia leads to parallel impairments in a number of different mechanisms supported by the affected cortices, such as AV temporal discrimination and reading related processes.

### 4.7.2 IMPLICIT TEMPORAL PROCESSING IS IMPAIRED IN DYSLEXIA

Overall, the ability to implicitly discriminate AV synchrony from asynchrony, represented by measures of McG SD and Stream-bounce SD, was poorer in the dyslexic sample relative to controls. This difference was reliably found across the TOJ and SJ non-speech dual-tasks. Within the AV speech tasks, the difference was significant in data from the SJ dual-task and borderline significant in the data from the TOJ task.

The window of AV integration was larger on average in the dyslexic groups only for AV speech stimuli. Dyslexic participants also showed higher rates of AV integration in the McGurk task, meaning that they had wider, flatter as well as taller AV integration functions. Why overall on average, the dyslexic group has wider and flatter profiles of AV integration in the speech tasks specifically could be related to overall higher rates of susceptibility to the McGurk illusion. Dyslexic participants might rely on visual information more so than controls in order to compensate for poor speech processing in the auditory modality (Ramirez & Mann, 2005; Savage et al., 2005; Ziegler et al., 2009).

Relying more on visual speech information to compensate for impairments in auditory speech processing might explain why dyslexic performance was poorer than controls in implicit discrimination of AV synchrony from asynchrony in the phoneme ID task. If participants overall relied on the lip-movements for the phoneme identification task, even at wide asynchronies, this would mean that their psychometric functions might begin further up on the y axis, at higher response proportions, leading to flatter functions and thus larger standard deviations. It cannot however also account for poorer dyslexic performance in implicit discrimination of AV synchrony in the non-speech integration task. Maximum levels of AV non-speech integration were comparable across groups in the SJ dual-task and lower in dyslexia in the TOJ task, and the widow of AV integration was equivalent across groups in both conditions. In other words, dyslexic individuals did not show greater levels of AV non-speech integration, nor did they show wider windows of AV integration, yet they did show poorer ability in implicit discrimination of AV synchrony from asynchrony in non-speech. Thus, the non-speech temporal processing impairment cannot be interpreted as a result of poor processing of the auditory stimulus in the same way that the AV speech temporal processing impairment can.

Overall higher rates of AV integration in the dyslexia sample might also account for why the difference between the two groups in implicit discrimination of AV synchrony is not as reliable in the speech data compared to the non-speech data; the former might be confounded by a greater tendency to rely on visual information in poor readers, which might in turn be caused by poorer auditory speech processing. This is proposed cautiously however, as the data does not allow for conjecture about auditory processing. An additional analysis revealed that across the sample, the ability to discriminate implicitly between AV synchrony and asynchrony was predictive of maximum levels of integration in both non-speech integration data sets. It is possible that poorer ability to discriminate between synchronous and asynchronous AV information might decrease the probability that visual and auditory information is integrated into and perceived as a unified percept. This might then hinder the formation of strong and automatic AV correspondences between phonemes and graphemes during early reading development, subsequently leading to less automaticity when retrieving auditory representations of written language and poorer reading fluency in adulthood. Poor implicit discrimination of AV synchrony might also impair the development of automaticity in reading by hindering the ability to selectively integrate written words and their auditory counterparts whilst inhibiting visual or auditory representations of words which succeed or precede the particular word being read, leading to less fluent reading, characteristic of hesitations, slower reading, letter reversals and word omissions.

Hairston et al.'s (2005) found that in non-speech AV integration, dyslexic individuals tended to integrate AV information at significantly larger AV asynchronies. Conversely here we found that the window of AV integration for non-speech stimuli was equivalent across the dyslexia and control groups and that the difference lies within implicit discrimination of AV synchrony from asynchrony. Hairston however only tested AV integration for auditory-lagging stimuli and not where the auditory led the visual. Thus, it is possible that if they had also tested temporal ventriloquism at auditory leads, on average the entire temporal window of integration may not have differed across the groups. The data in the current set were individually fitted with a function before group comparisons, whereas Hairston et al. used the raw data for group comparisons rather than fitting psychometric functions, thus the method of analysis also varied across this study and Hairston et al.'s. Sampling differences might also explain the different conclusions. Severity of reading impairments within dyslexia is subject to variability (Manis, Seidenberg, Doi, McBride-Chang, & Petersen, 1996) and the notion that individuals with dyslexia can be placed on a spectrum in terms of severity of reading impairments has been previously put forward (Rapcsak, Beeson, Henry, & Leyden, 2009). Hairston et al. did not specify the educational level of their sample and stated that all dyslexic performed below the 25th percentile on at least two decoding tasks. The individuals in the dyslexic sample here had been formally diagnosed with dyslexia by a certified institution and performed on average worse than controls. They were however students enrolled on a higher education degree meaning that the severity of dyslexia was probably not as high as it was in Hairston's sample.

In summary, implicit AV temporal discrimination seems to be impaired in dyslexia. Whereas the speech data were more difficult to interpret and could be confounded by auditory speech processing difficulties, there seems to be a reliable relationship between ability in AV temporal discrimination of non-speech and reading fluency.

### 4.7.3 AUDIOVISUAL INTEGRATION IN DYSLEXIA

Measures of the susceptibility to the McGurk effect from both SJ and TOJ dual-tasks showed that dyslexic individuals also tend to integrate incongruent AV speech information significantly more compared to controls. This finding is inconsistent with previous reports of unaffected AV McGurk integration in dyslexia (e.g. Bastien-

Toniazzo et al., 2010; Campbell et al., 1997). The inconsistent results between the current findings and previous studies on McGurk integration in dyslexia may be a result of the fact that the current experiments took into account individual differences in the AV asynchrony optimal for AV integration and measured AV integration as a function of AV asynchrony, whereas the previous studies presented AV only synchronously. The auditory component of the AV speech stimuli did contain some noise, in order to maximise the McGurk effect. The finding that on average, dyslexic individuals show higher levels of AV speech integration when the auditory speech is embedded in noise is also at odds with reports that dyslexic participants gain less benefit from visual information during AV speech-in-noise detection compared to typical readers (Ramirez & Mann, 2005). This could also be attributed to synchronous presentation of AV stimuli in Ramirez and Mann's study versus stimulus presentation at various AV asynchronies in the current study.

A higher degree of integration of incongruent AV information in dyslexia could reflect the possibility that auditory speech processes in dyslexia are impaired (Ramirez & Mann, 2005; Savage et al., 2005; Ziegler et al., 2009) and thus individuals might rely more on visual information when interpreting auditory speech. This interpretation is reflected in the observation that the window of AV speech integration and measures of implicit temporal discriminability in the dyslexia sample are also larger, suggesting that they also rely on visual information at wider AV asynchronies compared to typical readers. Because the dyslexic group did not also show higher levels of AV integration in the non-speech tasks supports this interpretation as well. Our finding that dyslexic individuals integrate AV speech stimuli more compared to controls is in line with a previous report that dyslexic individuals exhibit greater activation in multisensory integration areas such as the superior and middle temporal gyri, insula, basal ganglia, and the right inferior frontal and orbitofrontal cortex during perception of conflicting AV speech sounds which elicit the McGurk effect (Pekkola et al., 2006). Pekkola et al. reported that brain activity differed between groups despite no differences observed in the behavioural data. Their task however may have been too easy to detect any behavioural differences as the average accuracy of AV incongruence was 90% in fluent readers and 93% in dyslexic readers.

It is also possible that the higher rates of AV integration observed in the dyslexic sample could also be related to or accounted for by an impaired ability in detecting or processing AV incongruence. For example, Widmann et al. (2012) reported that dyslexic children show significantly poorer performance in discriminating between congruent and incongruent AV non-speech stimuli as well as differential brain activity to incongruent pairs of AV stimuli. In the same vein, Blau et al. (2009) reported that dyslexic individuals did not exhibit BOLD response suppression in multisensory areas comparable to that of controls when presented with incongruent letters and speech sounds. Furthermore, congruent AV information is more likely to be integrated (Spence, 2007). Dyslexic individuals may therefore be more likely to integrate incongruent AV speech information if they are less likely to notice incongruence between the visual and auditory modalities. Although there is no direct evidence for the relationship between AV integration and incongruence detection, factors such as semantic congruency (Doehrmann & Naumer, 2008) have been shown to increase the likelihood of AV integration. A higher rate of AV integration in dyslexic individuals could be a consequence of a poorer ability to detect incongruence in AV speech information. Such impairments could in turn lead to incorrect formation of AV correspondences between incongruent graphemes and phonemes or to a poorer ability to recognise incorrectly activated auditory representations of written letters or words during the development of reading. Examining the relationship between individual differences in the ability to detect congruence in AV stimuli and AV integration and how it may impact the formation of AV correspondences in reading development would be a relevant test for this hypothesis.

The finding that dyslexic individuals integrate incongruent AV information more so that controls is somewhat at odds with Froyen et al.'s (2011) conclusions that automatic AV integration is decreased in dyslexia. Froyen and colleagues based this conclusion on their findings that AV enhancement of early auditory MMN responses is non-existent in dyslexic children, whereas a later MMN response modulation was equivalent to that of typical readers. Whereas the sample here consisted of young adult readers, the sample tested by Froyen et al. consisted of 11 year old children, which according to the authors showed comparable brain activity to that of younger, less experienced but typical readers. It is therefore unclear whether the null effect of letters on the early auditory MMN in dyslexia reflects an AV integration impairment that is persistent into adulthood or whether it simply reflects delayed development of automatic AV integration in dyslexia. Visual modulation of auditory evoked potentials during processing of incongruent AV speech has been found to occur as early as 120-190ms post stimulus onset (Besle, Fort, Delpuech, & Giard, 2004). Incongruent lip-movements have also been reported to modulate auditory MMN responses occurring at 175ms (Saint-Amour, De Sanctis, Molholm, Ritter, & Foxe, 2007). Effects of incongruent visual information on auditory cortex activity seem to therefore occur well within the 100-250ms which within which the early MMN response was reported to occur in Foryen’s study. McGurk integration effects on auditory processing therefore do not seem to occur later than the letter-speech sound effect found to be absent in Froyen’s dyslexic sample. This means that the inconsistency between the behavioural results found here and the EEG results reported by Froyen cannot be accounted by a difference the latency at which visual information affects auditory processing across the two studies. A difference in the type of AV stimuli could however account it. Whereas Froyen used congruent letters and speech sounds, the current investigation used incongruent AV speech. There is a possibility that integration of congruent and incongruent AV information recruits different underlying mechanisms. For example incongruent AV integration has been reported to activate different neural structures (Benoit et al., 2010; Callan, Jones, Callan, & Akahane-Yamada, 2004; Sekiyama, 2003). It is therefore possible that AV integration of incongruent AV speech is intact but that the integration of congruent AV information is impaired.

Integration rates of AV non-speech stimuli in the Stream-Bounce illusion was found to be equivalent across the groups in the SJ dual-task but lower in the dyslexic group in TOJ dual-task. This difference could be due to additional attentional demands imposed by the TOJ task, which tends to overall be more difficult and essentially requires that the AV stimuli be segregated in order to make a judgement. Sluggish multimodal attention has been reported in dyslexic children (Facoetti et al., 2010; Facoetti, Lorusso, Cattaneo, Galli, & Molteni, 2005). It is possible that the dyslexic group were affected more by TOJ difficulty compared to controls and found it more difficult to switch attention to the Stream-Bounce task when performing the TOJ, leading to lower rates of AV integration in this dual-task. It is difficult for tasks difficulty to also account for the reverse effect seen in the McGurk tasks, where AV integration rates were higher for dyslexic individuals. The stimuli in the McGurk tasks however do not occur as quickly as does the collision point in the Stream-Bounce illusion, thus AV integration in the McGurk illusion may have withstood any attention related effects that may have been present in the Stream-bounce tasks.

### 4.7.4 CONCLUSION

In summary, both explicit and implicit AV temporal discrimination of speech seems to be impaired in dyslexia. Explicit temporal processing was related to reading ability, with poorer readers showing poorer AV temporal discrimination skills. Implicit temporal processing skills in dyslexia co-occurred with overall higher rates of AV speech integration in the group, suggesting that part of the poorer implicit temporal processing performance may have reflected an overall greater reliance on visual stimuli during the phone ID task. This is turn might have been a result of poor auditory processing, an interpretation which requires empirical testing. The results of the Stream-Bounce experiments indicate that explicit temporal processing of non-speech stimuli seems to be intact in dyslexia, as performance was comparable to that of controls. Analogous implicit measures of performance were however larger in the dyslexic group, suggesting that implicit temporal processing of non-speech stimuli is impaired, despite intact explicit temporal processing ability. Altogether these findings call for further research into the relationship between AV temporal processing and reading ability, which needs to consider that there may be a distinction between implicit and explicit temporal processing ability and that the two may be affected differentially in dyslexia across speech processing and non-speech processing.

### 4.7.5 REVISITING THE NEGATIVE CORRELATION FROM CHAPTER 2

In order to check the reliability of the negative correlation found in Chapter 2 between TOJ ePSS and iPSS, a correlation was run on the data between the two measures. A significant, negative correlation (illustrated in Figure 4.17) was found between McG iPSS and TOJ ePSS [*r*(33)=-.37, *p*=.036]. The correlation between iPSS and ePSS within the Stream-Bounce data did not reach significance [*r*(33)=.113, *p*=.618]. It seems that this antagonistic relationship is more reliable compared to that found in the Stream bounce illusion. The McGurk effect may be a more reliable way of measuring audiovisual integration, because it seems to provide more consistent results compared to Stream Bounce illusion and may be less prone to response bias.

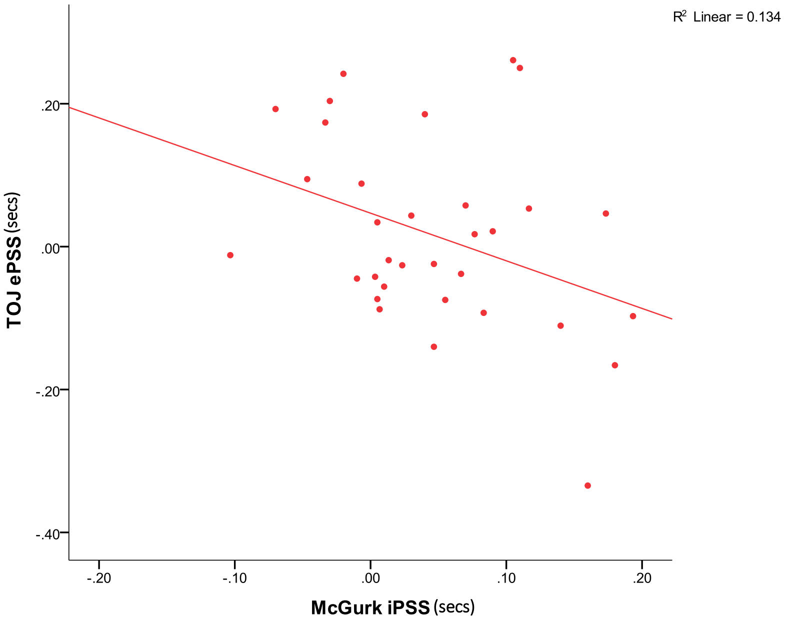


*Figure 4-17: Significant negative correlation between implicit PSS and explicit (TOJ) estimates of PSS*

# 5 CHAPTER 5: GENERAL DISCUSSION

## 5.1 OVERVIEW

Chapters 2 and 3 investigated whether audiovisual (AV) integration and explicit temporal judgements are supported by common or distinct underlying temporal mechanisms. In Chapter 2, behavioural correlations were carried out between estimates of subjective AV synchrony derived implicitly from AV integration tasks, and explicitly from temporal judgements. Chapter 3 examined whether individual differences in temporal sensitivity to AV asynchrony across implicit and explicit AV tasks are statistically dependent and whether they are related to structural variability in the same or different brain areas. The results of these chapters indicate that AV temporal processing may be supported by multiple task-specific mechanisms, whose performance relies on distinct neural substrates. Chapter 4 examined whether these potentially different temporal mechanisms might be impaired in dyslexia and whether individual differences in their performance are related to specific aspects of reading ability. Dyslexic individuals were found to show poorer sensitivity to AV speech asynchrony in both implicitly in AV speech integration and explicitly, in temporal judgements. This performance was linearly related to different aspects of reading ability. For non-speech temporal processing, dyslexic performance was found to be poorer only in the implicit AV task, and not in explicit temporal judgements. Performance in implicit non-speech temporal processing was related to aspects of reading ability.

The following chapter comprises of a summary of the findings of Chapters 2, 3 and 4. A discussion of the limitations of the experiments carried out in each chapter will follow each section. Where relevant, questions that remain unanswered and/or motivated by the findings in this thesis will be identified.

## 5.2 CHAPTER 2

Chapter 2 focused on the relationship between estimates of the point of subjective synchrony (PSS) derived from implicit and explicit temporal processing tasks. In other words, the chapter examined whether the asynchrony at which AV integration occurs maximally (iPSS) is also the asynchrony which maximally promotes the subjective perception of AV synchrony (ePSS) within the same individuals. Positive correlations between implicit and explicit estimates of the PSS would have been indicative of unity across the underlying temporal mechanisms of AV integration and explicit temporal judgements. On the contrary, significant negative correlations indicative of disunity were found between estimates of iPSS and estimates of TOJ ePSS derived across two different types of stimuli, and no significant relationships were found between estimates of iPSS and SJ ePSS. The negative correlation revealed between estimates derived from TOJ and the AV integration task proved difficult to explain using response bias and suggests disunity in subjective timing estimates across different multisensory processes.

The failure to find positive correlations between estimates of iPSS and ePSS suggests that AV integration and explicit temporal judgements might rely on different underlying temporal mechanisms which are subject to different neural delays. Individual subject analyses on the McGurk-SJ dual-task data of chapter 2 are also indicative of separate underlying timing mechanisms for SJ and AV integration. The logistic regressions carried out for each subject showed that for the majority of participants explicit perception of synchrony did not predict implicit perception of synchrony (i.e. audiovisual integration) over and above absolute AV asynchrony. The opposite was found for the Stream-Bounce data where for the majority, the explicit response did predict the implicit over and above AV asynchrony. This data was however subject to more types of response pattern, and the predominant response pattern was reflected in only less than 40% of the data sets. The Stream-Bounce task is also more prone to response bias, in that it is easy for subjects to use the bouncing percept to help them complete the timing task; if they were not sure whether the collision and beep were synchronous they may simply use bouncing percept as an indication of synchrony using higher level decision making. In light of this, the scenario that the data concurs mostly with seems to be that AV integration and explicit synchrony and temporal order perception are served by different underlying timing mechanisms, as shown in Figure 5.1 , below.

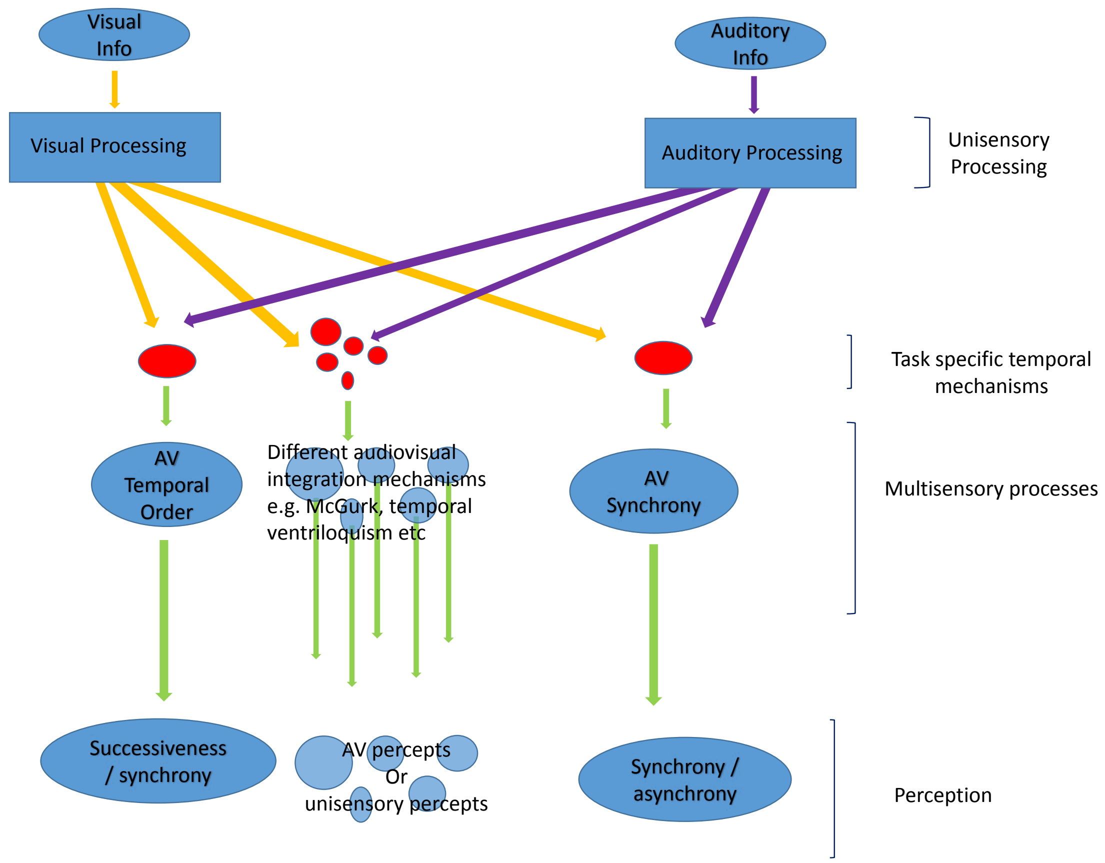


*Figure 5-1: Underlying cognitive processes contributing to SJs, TOJs and AV integration, with separate underlying temporal mechanisms for the different three tasks.*

Disunity of subjective AV timing across explicit temporal processing and AV integration is at odds which notions of centralised timing mechanisms which provide a common time code to multisensory events (Ivry & Schlerf, 2008; Ivry & Spencer, 2004; Nenadic et al., 2003). The results of Chapter 2 indicate that different multisensory processes might be subserved by task-dependent temporal mechanisms, subject to different information regarding the relative timing of single AV events. The extent to which individuals exhibit disunity across different AV processes as well as inter-individual variability in estimates of AV synchrony might result from differences in the structure and function of the brain across individuals. The negative correlation between iPSS and TOJ ePSS estimates is particularly difficult to explain if temporal recalibration (Fujisaki et al., 2004; Hanson et al., 2008; Vatakis et al., 2007) occurred across different multisensory processes; if this were the case, estimates of AV relative timing across different mechanisms would be recalibrated towards one another if they did not correspond with the observer's prior or present sensory experience and thus correlate positively. The renormalisation account however does not necessarily invalidate recalibration accounts as the latter might resynchronise estimates of AV timing within multisensory perceptual processes, rather than across. The antagonistic relationship suggests that estimates of AV relative timing derived from AV integration and temporal order judgements are being renormalised relative to an overall mean. The renormalisation account therefore proposes that AV synchrony estimates within localised, task-specific temporal mechanisms are assessed in relation to the average asynchrony across all temporal mechanisms, rather than recalibrated towards it.

### 5.2.1 IMPLICATIONS, LIMITATIONS AND FUTURE DEVELOPMENTS OF CHAPTER 2

The account of renormalisation raises some interesting testable hypotheses. The notion of multiple temporal mechanisms could be tested using sensory training methods. Powers et al. (2009) demonstrated that training on synchrony judgements with accuracy feedback for a period of 5 days led to a narrowing of the size of the temporal window of synchrony perception. Such effects should not generalise to the window of AV integration if explicit and implicit temporal processes are supported by distinct mechanisms. Improving performance in implicit temporal processing might however prove to be more challenging than improving performance at explicit temporal judgements.

Although Chapter 3 provided support for the notion that qualitatively different multisensory might be supported by distinct neural substrates, the distance between the clusters found to be related to performance in the different timing tasks was not necessarily large enough to account for some of the rather large discrepancies observed in the PSS data of Chapter 2. It is possible that sensory information travels via different routes to reach different multisensory processes, such as AV integration and temporal order processing. For example AV speech information in the McGurk effect might need to be processed by speech specific mechanisms before it arrives in AV integration areas to be unified, whereas this may not be necessary for a temporal order or a synchrony judgement to be performed. Such conjecture however goes beyond the present data and would need further investigation. One suitable method to test such a hypothesis might be Dynamic causal modelling (DCM) (Friston, Harrison, & Penny, 2003) using either functional MRI or EEG. This method examines the dynamic interaction between different cortical structures during the performance of a task by comparing predictive models of activity to brain activity observed. Critically, this method assumes that a change in stimulus or task demands can affect both the activity within a particular area, as well as the connectivity between different areas. DCM can be used to predict the different cortical pathways which might be activated as a result of an experimental manipulation and to infer causality between activity in one area on the activity of another (Friston et al., 2003). Investigations using such methodology could address whether there is a difference between the cortical networks which process auditory and visual information during AV integration and temporal judgements, which in turn might account for a difference in the relative arrival time of AV information at multisensory areas.

The renormalisation account (Freeman et al., 2013) could in theory also be tested using temporal recalibration (Fujisaki et al., 2004; Hanson et al., 2008; Vatakis et al., 2007), by adapting one mechanism to a particular AV asynchrony, and then measuring the PSS estimates from two mechanisms concurrently after exposure. Renormalisation would predict that if the PSS estimate from the adapted mechanism shifted in the direction of the asynchrony to which the observer was adapted, then the PSS from the unadapted mechanism should concurrently shift in the opposite direction as a result of being renormalised. Practically this would be challenging, due to the difficulty in controlling which mechanism is active at the time of exposure to asynchrony, or preventing a mechanisms to be active, which might result in integration mechanisms being adapted in some participants and explicit judgement mechanisms being adapted in others. One way in which such control could be implemented is by adapting participants to an asynchronous Stream-Bounce display (Sekuler et al., 1997) in which one disk is a different colour to the other, so that the two always appear to stream through each other. Consequently, participants would not integrate the auditory and visual information, and only mechanisms related to the temporal order would be recalibrated. After adaptation, implicit and explicit measures of PSS could be measured using the dual task paradigm used in this thesis. Whereas the TOJ PSS would be expected to shift towards the adapted asynchrony, the optimal asynchrony for bounce percepts, or in other words AV integration, should shift in the opposite direction.

The relationship between the PSS and structural variability was not examined in Chapter 3. One interpretation is that PSS measures are partially representative of the relative timing at which auditory and visual signals reach relevant cortices or multimodal areas as well as the time it takes for these signals to be processed (Keetels & Vroomen, 2012; King, 2005). These measures might therefore be related to a combination of structural density and structural connectivity. For example, PSS measures might be affected by processing speed within unisensory cortices, as well as the strength of connectivity between sensory receptors, unisensory cortices and multimodal areas. Furthermore, the sign and size of the PSS is likely to be dependent upon the strength of particular connections and the weakness of others, as well as upon other structural factors such as myelination which affect electrical transfer of signals in the brain. Our MRI scans did not convey such information, but the relationship between such factors and the estimates of the PSS is certainly an interesting topic for future developments, which we are already pursuing using improved technology such as Diffusion tensor imaging.

## 5.3 CHAPTER 3

Given that estimates of iPSS and ePSS Chapter 2 did not correlate positively and were indicative of multiple AV synchronisation mechanisms, Chapter 3 explored whether performance in AV temporal discrimination across implicit and explicit tasks is statistically dependent or independent and related to individual differences in the structure of the same or distinct anatomical brain areas. A strong positive relationship was found between the ability to discriminate AV synchrony from asynchrony across TOJs and SJs. No relationship was found between the ability to explicitly discriminate synchronous from asynchronous AV information in SJs and the ability to do so implicitly, during AV integration. Furthermore, the temporal windows of AV synchrony and integration were also statistically independent, and related to structural variability in different anatomical clusters. A moderate positive relationship was observed between discrimination ability in AV integration and TOJs, but only the latter was related to structural variability. Overall, higher sensitivity to AV synchrony in explicit AV timing, smaller temporal windows of AV synchrony and integration as well as higher rates AV integration were associated with increased density of grey matter volume in distinct subregions of the right temporal cortex. Performance in explicit temporal judgement tasks was related to grey matter volume in clusters located along the STS and overlapping with MTG and STG, whereas the size of the temporal window of AV integration was related to a cluster located more posteriorly, within the MTG. The results of Chapter 3 indicate that AV relative timing is processed via different task-dependent mechanisms, whose performance is likely be supported by distinct neuronal populations. The different parameters extracted from the temporal profile of AV integration were in turn related to distinct areas of the right temporal cortex, suggesting that complementary processes of AV integration are supported by different neural substrates.

### 5.3.1 IMPLICATIONS LIMITATIONS AND FUTURE DEVELOPMENTS OF CHAPTER 3

Chapter 3 showed a relationship between performance in AV timing and integration ability and the structure of the brain. However, the analysis methodology used here is correlational, so in order to test the hypothesis that the structures revealed in this investigation might support the functional role of the processes measured here behaviourally, a more direct causal relationship would need to be established. Brain stimulation and disruption techniques such as transcranial direct current stimulation (tDCS) and transcranial magnetic stimulation (TMS) are well suited complementary methods for correlational investigations of function, structure and behaviour. Disruptive TMS for example has been successfully used alongside structural (Kanai et al., 2010; Kanai, Carmel, Bahrami, & Rees, 2011) and functional imaging (Beauchamp et al., 2010) to confirm the functional role of brain structures whose activity or structure was related to individual variability in perception. Here, it would be predicted that applying repetitive TMS (rTMS) to the coordinates identified in Chapter 3 should impair performance at the different tasks each of the areas are related to, specifically by altering the specific parameters of the temporal functions which describe task performance. Moreover, using TMS could also confirm whether indeed the different tasks used in this investigation rely on independent mechanisms. For example, applying rTMS to the cluster associated with TOJ performance should decrease performance in temporal order judgements, but have no detrimental effect on implicit temporal processing during AV integration if the underlying temporal and neural mechanisms of these tasks are indeed distinct as suggested by this thesis.

There is also a possibility that structural and functional connectivity could explain variability in behavioural measures that was not explained by grey matter volume. For example, despite a strong positive relationship between the SD parameters from SJs and TOJs, the two measures were related to different clusters of grey matter volume. This indicates the two different clusters explain variability in their corresponding performance measure which is not shared with the variability in the other measure. On the other hand, the shared variability across TOJ and SJ discrimination measures demonstrated by the positive correlation between them might be determined by other factors, such as the consistency in the time taken of auditory and visual to reach multisensory cortices after entering the nervous system which may depend on other structural or even functional variability, such as the strength of structural and functional connectivity in the nervous system.

It is difficult to infer whether individual variability in grey matter volume arises as a result of cortical development or due to neuronal plasticity during adulthood. Variability in brain structure observed between individuals is likely to be in part determined by genetic factors (Thompson et al., 2001), but it is likely to also result from behavioural experience during the lifetime (Sale, Berardi, & Maffei, 2012). For example in the motor domain, there is evidence that prolonged training in tasks can lead to increases in grey matter density in areas associated with the performance of those tasks in children (Hyde et al., 2009), younger adults (Draganski, Gaser, & Busch, 2004) as well as older adults (Boyke, Driemeyer, Gaser, Büchel, & May, 2008). However, such training induced structural changes seem be temporary and to disappear if training is discontinued (Draganski & May, 2008), and whether similar structural plasticity can be induced by training in areas associated with perceptual and higher cognitive process is still to be determined (Kanai & Rees, 2011). Nevertheless, studies which investigate the link between structure and behaviour such as the one described in this thesis and investigations into structural plasticity such as those cited here have major implications for developmental disorders in which structural brain differences are observed, such as autism (Bahrick & Todd, 2010; Bebko, Weiss, Demark, & Gomez, 2006; Mongillo et al., 2008; Taylor, Isaac, & Milne, 2010) and dyslexia (Eckert et al., 2005; Eckert, 2004).

## 5.4 CHAPTER 4

Chapter 4 examined performance in AV integration and in implicit and explicit temporal processing in dyslexic and typical readers. AV temporal processing ability in the context of both AV integration and explicit temporal judgements was compared between groups. Reading ability was not related to AV integration and timing, over and above dyslexia. Explicit and implicit AV temporal discrimination of speech was found to be less sensitive in dyslexia. Overall, dyslexic readers showed less sensitivity in discriminating temporal order and synchrony suggesting AV temporal discrimination and reading ability. Dyslexic individuals overall showed higher rates of AV speech integration, or in other words a greater reliance on visual stimuli during the phone ID task. This could be a result of poor auditory processing, and may have contributed somewhat to poorer temporal discrimination ability. Explicit temporal processing of non-speech stimuli seems to be intact in dyslexia, as the group's performance in the non-speech explicit temporal judgements was comparable to that of controls. Dyslexic performance in implicit synchrony discrimination of non-speech was however poorer. In non-speech AV processing, implicit but not explicit temporal processing is therefore impaired in dyslexia. Performance in implicit temporal processing was related to reading ability across the entire sample, most reliably across the non-speech tasks, compared to the speech tasks.

### 5.4.1 IMPLICATIONS LIMITATIONS AND FUTURE DEVELOPMENTS OF CHAPTER 4

Chapter 4 of this thesis found that performance in temporal processing of AV speech in dyslexia seems to be impaired compared to undiagnosed controls and that it is linearly related to reading ability. Chapter 3 found that increased grey matter volume in clusters located in superior temporal gyrus (STG) was related to better performance at the same tasks performed by participants the AV speech experiments of Chapter 2. The STG happens to be one of the brain areas reported to be subject to structural differences in Dyslexia, including lower levels of grey matter density (Brown, Eliez, Menon, Rumsey, White, Reiss, 2001; Eckert et al., 2008). If impairments in AV temporal processing in dyslexia are a result of developmental structural abnormalities, and training can modify brain structure by increasing grey matter volume, then both the structure of the neural mechanisms underlying AV temporal processing and behavioural performance might be improved with behavioural training in AV temporal tasks. Perceptual training has been shown in previous literature to improve performance in AV temporal judgements (Powers, Hillock, & Wallace, 2009). Chapter 4 showed not only that temporal processing performance was poorer in dyslexia, but that it was also linearly related to reading ability across all participants. An intriguing area for future research, and a good test of whether AV temporal processing can be attributed a causal role in reading ability is examining whether the effects of behavioural training in AV temporal tasks such as the ones reported by Powers et al., (2009) might also ameliorate reading impairments associated with poor AV temporal processing.

However, before such hypothesis could be tested, there are some issues related to the studies on chapter 4 that would need to be addressed empirically. As discussed in section 4.7, there is a possibility that at least part of the difference between the two groups found in the ability to discriminate AV synchrony might be explained by poor auditory (unisensory) processing, because dyslexic readers on average showed a higher tendency to rely on lip-movements in the phoneme identification task. This possibility should be ruled out in future research by measuring and controlling for performance in unisensory speech detection tasks. There is also a possibility that unisensory timing impairments could be responsible for the multisensory timing deficit found in chapter 4, although it is unclear why implicit and not explicit timing would be affected if this were the case. Further research is needed to explore this possibility. Furthermore, participants from both the dyslexic and control group had to be excluded due to not being susceptible to the AV illusions used to measure implicit AV timing.

Participants in both the control and dyslexic group in chapter 4 were undergraduate students at university. The samples were assumed to have similar levels of IQ as they originated from a sample with similar educational background (see Bright, Jaldow, & Kopelman, 2002 for demographic information such as educational level as a predictor of intelligence), there is a possibility that the IQ levels across the two groups may not have been equivalent. IQ has been reported in the past to be related to temporal processing ability (Rammsayer & Brandler, 2007). If the dyslexic sample did have a lower IQ, this may explain why their temporal processing ability was poorer in the current tasks. However, there is also a possibility that those in the dyslexic group had higher IQs compared to controls, suggested by the fact they have achieved the entry criteria for (the same) university, despite the reading difficulties which may have made their learning more difficult. If this were the case, a difference in IQ would not explain the difference in temporal processing ability found between the groups.

Throughout this entire thesis, AV integration was measured concurrently with explicit timing judgements. In real life, it is likely that several processes such as AV integration and explicit timing will be active and used at the same time, so the dual-task context in which performance in explicit and implicit temporal processing is not entirely artificial. However, the possibility that having to perform two tasks at the same time may have contributed to some of the differences found between dyslexic and typical readers, for example by increasing attentional load, which might mean that poor temporal processing is not directly linked to reading, but that the relationships and differences found may be mediated by attentional factors. Thus, in order to exclude this possibility a replication of the findings in a single task setting is part of our ongoing plan for further research.

As the review on dyslexia in Chapter 4 demonstrated, dyslexia is complex, multi-dimensional disorder, associated with a range of sensory processing impairments. An occurring finding of the literature reviewed is that sensory processing impairments are not typical of all dyslexic individuals, and often impairments are only found in subgroups of individuals tested. Thus, one possibility is that the condition comprises subtypes which are likely to be characterised not only by the type of reading impairments individuals exhibit, but also by their performance at sensory and temporal processing tasks, which might underlie individual variability in qualitatively different aspects of reading ability. One way in which Chapter 4 contributed to the existing literature was by presenting findings indicative of multisensory temporal processing impairments in dyslexia, on which the literature is scarce. One of the remaining questions is how multisensory impairments might covary with unisensory deficits, as well as with specific aspects of reading ability in dyslexia as well as in the general population. A large scale investigation into profiling unisensory and multisensory processing ability, as well as specific aspects of reading ability in dyslexic and typical populations would help answer this. For example, cluster analysis on data from a large sample would be a suitable method for addressing this question and for subtyping dyslexia according to reading ability together with performance at basic perceptual processing.

Determining which aspects of sensory and temporal processing, both within and across modalities are present within dyslexic individuals would further our understanding of the disorder and might also lead to better and earlier diagnosis. For example, diagnosis of dyslexia can only be made once an individual has begun to learn to read if the diagnostic criteria are based on reading ability. In contrast, sensory and temporal processing ability can be measured and assessed much earlier than reading ability and could prove to be a useful tool in diagnosing dyslexia earlier in development. Understanding what the different subtypes of dyslexia are and which aspects of sensory processing are related to reading impairment might also lead to impairment-specific remediation methods, tailored to the individual.

To conclude, Chapter 4 has shown that AV temporal processing is impaired in dyslexia. Further research is needed to determine how these impairments might relate to other sensory processing problems known to exist in dyslexia as well as how together, they relate to specific aspect of reading ability. A better understanding of subtypes of dyslexia based on aspects of reading ability as well as unisensory and multisensory processing could lead to developments in diagnosis, remediation and education in dyslexia.

## 5.5 SUMMARY

In summary, this thesis examined individual differences in AV temporal processing across qualitatively different tasks, which require that the temporal relation between auditory and visual information is processed either implicitly or explicitly. Correlations between measures of performance across implicit and explicit temporal processing tasks within individuals and correlations between these measures and individual differences in brain structure yielded results which indicate that AV relative timing might be processed by multiple task-specific temporal mechanisms, whose performance is supported by different neuronal populations. Furthermore, the thesis revealed that the ability to discriminate between synchronous and asynchronous AV information is poorer in dyslexia across both implicit and explicit temporal processing of AV speech and in implicit temporal processing of AV non-speech information. Performance that was found to be poorer in dyslexic individuals was also found to be linearly related to reading ability across the entire sample of controls and dyslexic readers. Altogether, the thesis has contributes new knowledge about individual differences, multisensory interactions, brain anatomy and dyslexia. The results of this thesis motivate interesting testable hypotheses and future developments about the underlying mechanisms of the subjective perception of AV relative timing, about the structural and functional mechanisms which might support them. Furthermore, the thesis demonstrated that AV temporal processing is related to reading ability, emphasising the importance of further research into AV processes in the context of higher cognitive functions.

# REFERENCES

Adhikari, B. M., Goshorn, E. S., Lamichhane, B., & Dhamala, M. (2013). Temporal-order judgment of audiovisual events involves network activity between parietal and prefrontal cortices. *Brain Connectivity*, *3*(5), 536–45. doi:10.1089/brain.2013.0163

Advanced Television Systems Committee. (2003). ATSC Implementation Subcommittee Finding : Relative Timing of Sound and Vision for Broadcast, (June).

Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, *15*(9), 839–43. doi:10.1016/j.cub.2005.03.046

Amitay, S., Ahissar, M., & Nelken, I. (2002). Auditory processing deficits in reading disabled adults. *Journal of the Association for Research in Otolaryngology : JARO*, *3*(3), 302–20. doi:10.1007/s101620010093

Amitay, S., Ben-Yehudah, G., Banai, K., & Ahissar, M. (2002). Disabled readers suffer from visual and auditory impairments but not from a specific magnocellular deficit. *Brain : A Journal of Neurology*, *125*(Pt 10), 2272–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12244084

Asakawa, K. (2008). Temporal Recalibration in Audio-Visual Speech Integration Using a Simultaneity Judgment Task and the McGurk Identification Task. *Electrical Communication*, *1*, 1669–1673.

Aschersleben, G., & Bertelson, P. (2003). Temporal ventriloquism : crossmodal interaction on the time dimension 2. Evidence from sensorimotor synchronization. *International Journal of Psychophysiology*, *50*, 157–163. doi:10.1016/S0167-8760

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews. Neuroscience*, *7*(5), 358–66. doi:10.1038/nrn1888

Bahrick, L., & Todd, J. (2010). Impaired detection of temporal synchrony for social and nonsocial events in children with autism spectrum disorders. *International Meeting for Autism Research, Philadelphia*, *50*. Retrieved from http://infantlab.fiu.edu/Posters/IMFAR 2010 ASD Intersensory Poster Layout.pdf

Bastien-Toniazzo, M., Stroumza, A., & Cavé, C. (2010). Audio-Visual Perception and Integration in Developmental Dyslexia: An Exploratory Study Using the McGurk

Effect. *Current Psychology Letters. Behaviour, Brain & Cognition*, *25*(3). Retrieved from http://cpl.revues.org/index4928.html

Baum, S. H., Martin, R. C., Hamilton, a C., & Beauchamp, M. S. (2012). Multisensory speech perception without the left superior temporal sulcus. *NeuroImage*, *62*(3), 1825–1832. doi:10.1016/j.neuroimage.2012.05.034

Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, *41*(5), 809–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15003179

Beauchamp, M. S., Nath, A. R., & Pasalar, S. (2010). fMRI-Guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(7), 2414–7. doi:10.1523/JNEUROSCI.4865-09.2010

Bebko, J. M., Weiss, J. a, Demark, J. L., & Gomez, P. (2006). Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*(1), 88–98. doi:10.1111/j.1469-7610.2005.01443.x

Benedetti, F., Bernasconi, A., Bosia, M., Cavallaro, R., Dallaspezia, S., Falini, A., … Smeraldi, E. (2009). Functional and structural brain correlates of theory of mind and empathy deficits in schizophrenia. *Schizophrenia Research*, *114*(1-3), 154–60. doi:10.1016/j.schres.2009.06.021

Benoit, M. M., Raij, T., Lin, F.-H., Jääskeläinen, I. P., & Stufflebeam, S. (2010). Primary and multisensory cortical activity is correlated with audiovisual percepts. *Human Brain Mapping*, *31*(4), 526–38. doi:10.1002/hbm.20884

Bertelson, P. (2003). Temporal ventriloquism: crossmodal interaction on the time dimension 1. Evidence from auditory–visual temporal order judgment. *International Journal of Psychophysiology*, *50*(1-2), 147–155. doi:10.1016/S0167-8760(03)00130-2

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & …*, *62*(2), 321–332. Retrieved from http://link.springer.com/article/10.3758/BF03205552

Bertini, C., Leo, F., Avenanti, A., & Làdavas, E. (2010). Independent mechanisms for ventriloquism and multisensory integration as revealed by theta-burst stimulation. *The European Journal of Neuroscience*, *31*(10), 1791–9. doi:10.1111/j.1460-9568.2010.07200.x

Besle, J., Fort, A., Delpuech, C., & Giard, M.-H. (2004). Bimodal speech: early suppressive visual effects in human auditory cortex. *The European Journal of Neuroscience*, *20*(8), 2225–34. doi:10.1111/j.1460-9568.2004.03670.x

Bickart, K. C., Wright, C. I., Dautoff, R. J., Dickerson, B. C., & Barrett, L. F. (2011). Amygdala volume and social network size in humans. *Nature Neuroscience*, *14*(2), 163–4. doi:10.1038/nn.2724

Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., & Blomert, L. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: a functional magnetic resonance imaging study of dyslexic children. *Brain : A Journal of Neurology*, *133*(Pt 3), 868–79. doi:10.1093/brain/awp308

Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Current Biology : CB*, *19*(6), 503–8. doi:10.1016/j.cub.2009.01.065

Blomert, L., & Froyen, D. (2010). Multi-sensory learning and learning to read. *International Journal of Psychophysiology*, *77*(3), 195–204. doi:10.1016/j.ijpsycho.2010.06.025

Boddaert, N., Chabane, N., Gervais, H., Good, C. D., Bourgeois, M., Plumet, M.-H., … Zilbovicius, M. (2004). Superior temporal sulcus anatomical abnormalities in childhood autism: a voxel-based morphometry MRI study. *NeuroImage*, *23*(1), 364–9. doi:10.1016/j.neuroimage.2004.06.016

Boenke, L. T., Deliano, M., & Ohl, F. W. (2009). Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *198*(2-3), 233–44. doi:10.1007/s00221-009-1917-z

Boets, B., Wouters, J., van Wieringen, A., & Ghesquière, P. (2007). Auditory processing, speech perception and phonological ability in pre-school children at high-risk for dyslexia: a longitudinal study of the auditory temporal processing theory. *Neuropsychologia*, *45*(8), 1608–20. doi:10.1016/j.neuropsychologia.2007.01.009

Boyke, J., Driemeyer, J., Gaser, C., Büchel, C., & May, A. (2008). Training-induced brain structure changes in the elderly. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(28), 7031–5. doi:10.1523/JNEUROSCI.0742-08.2008

Breznitz, Z., & Meyler, A. (2003). Speed of lower-level auditory and visual processing as a basic factor in dyslexia: Electrophysiological evidence. *Brain and Language*, *85*(2), 166–184. doi:10.1016/S0093-934X(02)00513-8

Bright, P., Jaldow, E., & Kopelman, M. D. (2002). The National Adult Reading Test as a measure of premorbid intelligence: A comparison with estimates derived from demographic variables. *Journal of the International Neuropsychological Society*, *8*(6), 847–854. doi:10.1017/S1355617702860131

Brown, L. N., & Sainsbury, R. S. (2002). Age- and sex-related differences in temporal judgments to visual stimuli: support for hemispheric equivalence. *Perception & Psychophysics*, *64*(5), 693–702. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12201329

Brown, W. E., Eliez, S., Menon, V., Rumsey J. M. , White, C. D., Reiss, A. L. (2001). Preliminary evidence of widespread morphological variations of the brain in dyslexia. *Neurology*, *56*, 781–783. Retrieved from about:blank

Callan, D. E., Jones, J. a, Callan, A. M., & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *NeuroImage*, *22*(3), 1182–94. doi:10.1016/j.neuroimage.2004.03.006

Callan, D. E., Jones, J. a, Munhall, K., Callan, A. M., Kroos, C., & Vatikiotis-Bateson, E. (2003). Neural processes underlying perceptual enhancement by visual speech gestures. *Neuroreport*, *14*(17), 2213–8. doi:10.1097/01.wnr.0000095492.38740.8f

Calvert, G. a, Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology : CB*, *10*(11), 649–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10837246

Calvert, G. a, Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. *NeuroImage*, *14*(2), 427–38. doi:10.1006/nimg.2001.0812

Campbell, R., Whittingham, A., Frith, U., Massaro, D. W., & Cohen, M. (1997). Audiovisual speech perception in dyslexics: Impaired unimodal perception but no audiovisual integration deficit. In *Audio-Visual Speech Processing: Computational & Cognitive Science Approaches* (pp. 85–88). Retrieved from http://www.isca-speech.org/archive_open/avsp97/av97_085.html

Campbell, R., Whittingham, A., Frith, U., Massaro, D. W., Cohen, M. M., Massaro, D. W., & Cohen, M. (1997). Audiovisual speech perception in dyslexics: Impaired unimodal perception but no audiovisual integration deficit. In *Audio-Visual Speech Processing: Computational & Cognitive Science Approaches* (pp. 85–88). Retrieved from http://www.isca-speech.org/archive_open/avsp97/av97_085.html

Chase, C., & Jenner, A. R. (1993). Magnocellular visual deficits affect temporal processing of dyslexics. *Annals of the New York Academy of Sciences*, *682*(1), 326–329. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.1993.tb22983.x/abstract

Chung, M. K., Dalton, K. M., Alexander, A. L., & Davidson, R. J. (2004). Less white matter concentration in autism: 2D voxel-based morphometry. *NeuroImage*, *23*(1), 242–51. doi:10.1016/j.neuroimage.2004.04.037

Cook, L. a, Van Valkenburg, D. L., & Badcock, D. R. (2011). Predictability affects the perception of audiovisual synchrony in complex sequences. *Attention, Perception & Psychophysics*, *73*(7), 2286–97. doi:10.3758/s13414-011-0185-8

Cornelissen, P., Richardson, a, Mason, a, Fowler, S., & Stein, J. (1995). Contrast sensitivity and coherent motion detection measured at photopic luminance levels in dyslexics and controls. *Vision Research*, *35*(10), 1483–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7645277

Davis, B., Christie, J., & Rorden, C. (2009). Temporal order judgments activate temporal parietal junction. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(10), 3182–8. doi:10.1523/JNEUROSCI.5793-08.2009

De Martino, S., Espesser, R., Rey, V., & Habib, M. (1999). The “temporal processing deficit” hypothesis in dyslexia: new experimental evidence. *Brain and Cognition*, *46*(1-2), 104–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11527306

DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., & Gray, J. R. (2010). Testing predictions from personality neuroscience. Brain structure and the big five. *Psychological Science*, *21*(6), 820–8. doi:10.1177/0956797610370159

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, *9*(6), 719–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7220244

Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: how meaning modulates processes of audio-visual integration. *Brain Research*, *1242*, 136–50. doi:10.1016/j.brainres.2008.03.071

Dole, M., Meunier, F., & Hoen, M. (2013). Gray and White Matter Distribution in Dyslexia: A VBM Study of Superior Temporal Gyrus Asymmetry. *PloS One*, *8*(10), e76823. doi:10.1371/journal.pone.0076823

Draganski, B., Gaser, C., & Busch, V. (2004). Neuroplasticity: changes in grey matter induced by training. *Nature*. Retrieved from http://www.nature.com/nature/journal/v427/n6972/abs/427311a.html

Draganski, B., & May, a. (2008). Training-induced structural changes in the adult human brain. *Behavioural Brain Research*, *192*(1), 137–42. doi:10.1016/j.bbr.2008.02.015

Eagleman, D. M. (2008). Human time perception and its illusions. *Current Opinion in Neurobiology*, *18*(2), 131–6. doi:10.1016/j.conb.2008.06.002

Eckert, M. (2004). Neuroanatomical markers for dyslexia: a review of dyslexia structural imaging studies. *The Neuroscientist : A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *10*(4), 362–71. doi:10.1177/1073858404263596

Eckert, M. a, Lombardino, L. J., Walczak, A. R., Bonihla, L., Leonard, C. M., & Binder, J. R. (2008). Manual and automated measures of superior temporal gyrus asymmetry: concordant structural predictors of verbal ability in children. *NeuroImage*, *41*(3), 813–22. doi:10.1016/j.neuroimage.2008.03.002

Eckert, M., Leonard, C., Wilke, M., Richards, T., Richards, A., & Berninger, V. (2005). Anatomical Signatures of Dyslexia in Children: Unique Information from Manual and Voxel Based Morphometry Brain Measures. *Cortex*, *41*(3), 304–315. doi:10.1016/S0010-9452(08)70268-5

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, *99*(1), 3–19. doi:10.1037//0033-2909.99.1.3

Facoetti, A., Lorusso, M. L., Cattaneo, C., Galli, R., & Molteni, M. (2005). Visual and auditory attentional capture are both sluggish in children with developmental dyslexia. *Acta Neurobiologiae Experimentalis*, *65*(1), 61–72. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15794032

Facoetti, A., Trussardi, A. N., Ruffino, M., Lorusso, M. L., Cattaneo, C., Galli, R., … Zorzi, M. (2010). Multisensory spatial attention deficits are predictive of phonological decoding skills in developmental dyslexia. *Journal of Cognitive Neuroscience*, *22*(5), 1011–25. doi:10.1162/jocn.2009.21232

Farmer, M. E., & Klein, R. (1993). Auditory and visual temporal processing in dyslexic and normal readers. *Annals of the New York Academy of Sciences*, *682*(1), 339–341. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.1993.tb22987.x/pdf

Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin & Review*, *2*(4), 460–493. Retrieved from http://www.springerlink.com/index/V17513250078RJ69.pdf

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science (New York, N.Y.)*, *329*(5998), 1541–3. Retrieved from http://www.sciencemag.org/content/329/5998/1541.abstract

Freeman, E. D., Ipser, A., Palmbaha, A., Paunoiu, D., Brown, P., Lambert, C., … Driver, J. (2013). Sight and sound out of synch: Fragmentation and renormalisation of audiovisual integration and subjective timing. *Cortex*, 1–13. doi:10.1016/j.cortex.2013.03.006

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273–1302. doi:10.1016/S1053-8119(03)00202-7

Froyen, D., van Atteveldt, N., & Blomert, L. (2010). Exploring the Role of Low Level Visual Processing in Letter-Speech Sound Integration: A Visual MMN Study. *Frontiers in Integrative Neuroscience*, *4*(April), 9. doi:10.3389/fnint.2010.00009

Froyen, D., Van Atteveldt, N., Bonte, M., & Blomert, L. (2008). Cross-modal enhancement of the MMN to speech-sounds indicates early and automatic integration of letters and speech-sounds. *Neuroscience Letters*, *430*(1), 23–8. doi:10.1016/j.neulet.2007.10.014

Froyen, D., Willems, G., & Blomert, L. (2011, July). Evidence for a specific cross-modal association deficit in dyslexia: an electrophysiological study of letter-speech sound processing. *Developmental Science*. doi:10.1111/j.1467-7687.2010.01007.x

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, *7*(7), 773–8. doi:10.1038/nn1268

Funnell, M. G., Corballis, P. M., & Gazzaniga, M. S. (2003). Temporal discrimination in the split brain. *Brain and Cognition*, *53*(2), 218–222. doi:10.1016/S0278-2626(03)00113-1

García-Pérez, M. a, & Alcalá-Quintana, R. (2012). On the discrepant results in synchrony judgment and temporal-order judgment tasks: a quantitative model. *Psychonomic Bulletin & Review*, *19*(5), 820–46. doi:10.3758/s13423-012-0278-y

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, *120*(3), 453–63. doi:10.1016/j.clinph.2008.11.029

Ghazanfar, A. a, & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*(6), 278–85. doi:10.1016/j.tics.2006.04.008

Gogate, L. J., & Bahrick, L. E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology*, *69*(2), 133–49. doi:10.1006/jecp.1998.2438

Goswami, U. (2000). Phonological representations, reading development and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia*, *151*, 133–151. Retrieved from http://www.ingentaconnect.com/content/jws/dys/2000/00000006/00000002/art00160

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, *72*(3), 561–582. doi:10.3758/APP

Guski, R., & Troje, N. F. (2003). Audiovisual phenomenal causality. *Perception & Psychophysics*, *65*(5), 789–800. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12956586

Hairston, W. D., Burdette, J. H., Flowers, D. L., Wood, F. B., & Wallace, M. T. (2005). Altered temporal profile of visual-auditory multisensory interactions in dyslexia. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *166*(3-4), 474–80. doi:10.1007/s00221-005-2387-6

Halverson, H. E., & Freeman, J. H. (2010). Medial auditory thalamic input to the lateral pontine nuclei is necessary for auditory eyeblink conditioning. *Neurobiology of Learning and Memory*, *93*(1), 92–8. doi:10.1016/j.nlm.2009.08.008

Hämäläinen, J. a, Salminen, H. K., & Leppänen, P. H. T. (2013). Basic auditory processing deficits in dyslexia: systematic review of the behavioral and event-related potential/ field evidence. *Journal of Learning Disabilities*, *46*(5), 413–27. doi:10.1177/0022219411436213

Hanson, J. V. M., Heron, J., & Whitaker, D. (2008). Recalibration of perceived time across sensory modalities. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *185*(2), 347–52. doi:10.1007/s00221-008-1282-3

Harris, L. R., Harrar, V., Jaekl, P., & Kopinska, A. (2008). Mechanisms of simultaneity constancy. In R. Nijhawan (Ed.), *Issues of space and time in perception and action*. Cambridge University Press. Retrieved from http://scholar.google.com/scholar?start=100&q=simultaneity+perception&hl=en#1

Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., & Shakespeare, B. (2009). Speech perception abilities of adults with dyslexia: is there any evidence for a true deficit? *Journal of Speech, Language, and Hearing Research : JSLHR*, *52*(6), 1510–29. doi:10.1044/1092-4388(2009/08-0220)

Heron, J., Hanson, J., & Whitaker, D. (2007). Asynchrony adaptation is not specific to sound and vision. *Ophthalmic and Physiological Optics*, *28*(1), 4–5. doi:10.1111/j.1475-1313.2007.00530.x

Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*, *69*(2), 387–96. doi:10.1016/j.neuron.2010.12.027

Hollich, G., Newman, R. S., & Jusczyk, P. W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development*, *76*(3), 598–613. doi:10.1111/j.1467-8624.2005.00866.x

Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(31), 13022–7. doi:10.1073/pnas.0901123106

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical training shapes structural brain development. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(10), 3019–25. doi:10.1523/JNEUROSCI.5118-08.2009

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences*, *12*(7), 273–80. doi:10.1016/j.tics.2008.04.002

Ivry, R. B., & Spencer, R. M. C. (2004). The neural representation of time. *Current Opinion in Neurobiology*, *14*(2), 225–32. doi:10.1016/j.conb.2004.03.013

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *NeuroImage*, *20*(2), 693–712. doi:10.1016/S1053-8119(03)00343-4

Johnston, A., Bruno, A., Watanabe, J., Quansah, B., Patel, N., Dakin, S., & Nishida, S. (2008). Visually-based temporal distortion in dyslexia. *Vision Research*, *48*(17), 1852–8. doi:10.1016/j.visres.2008.04.029

Jones, J. a, & Callan, D. E. (2003). Brain activity during audiovisual speech perception: an fMRI study of the McGurk effect. *Neuroreport*, *14*(8), 1129–33. doi:10.1097/01.wnr.0000074343.81633.2a

Kanai, R., Bahrami, B., & Rees, G. (2010). Human parietal cortex structure predicts individual differences in perceptual rivalry. *Current Biology : CB*, *20*(18), 1626–30. doi:10.1016/j.cub.2010.07.027

Kanai, R., Carmel, D., Bahrami, B., & Rees, G. (2011). Structural and functional fractionation of right superior parietal cortex in bistable perception. *Current Biology*, *21*(3), R106–7. doi:10.1016/j.cub.2010.12.009

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*, *12*(4), 231–42. doi:10.1038/nrn3000

Kanai, R., Sheth, B. R., Verstraten, F. a J., & Shimojo, S. (2007). Dynamic perceptual changes in audiovisual simultaneity. *PloS One*, *2*(12), e1253. doi:10.1371/journal.pone.0001253

Keetels, M., & Vroomen, J. (2005). The role of spatial disparity and hemifields in audio-visual temporal order judgments. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *167*(4), 635–40. doi:10.1007/s00221-005-0067-1

Keetels, M., & Vroomen, J. (2007). No effect of auditory-visual spatial disparity on temporal recalibration. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *182*(4), 559–65. doi:10.1007/s00221-007-1012-2

Keetels, M., & Vroomen, J. (2012). Perception of Synchrony between the Senses. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes* (pp. 1–27). Boca Raton (FL): CRC Press.

Kim, P., Leckman, J. F., Mayes, L. C., Newman, M.-A., Feldman, R., & Swain, J. E. (2010). Perceived quality of maternal care in childhood and structure and function of mothers' brain. *Developmental Science*, *13*(4), 662–73. doi:10.1111/j.1467-7687.2009.00923.x

King, A. J. (2005). Multisensory Integration: Strategies for Synchronization. *Current Biology*, *15*(9), R336–9. doi:10.1016/j.cub.2005.04.025

Kinsbourne, M., Rufo, D. T., Gamzu, E., Palmer, R. L., & Berliner, A. K. (1991). NEUROPSYCHOLOGICAL DEFICITS IN ADULTS WITH DYSLEXIA. *Developmental Medicine & Child Neurology*, *33*(9), 763–775. doi:10.1111/j.1469-8749.1991.tb14960.x

Kolomiets, B. P., Deniau, J. M., Mailly, P., Ménétrey, a, Glowinski, J., & Thierry, a M. (2001). Segregation and convergence of information flow through the cortico-subthalamic pathways. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(15), 5764–72. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11466448

Kronbichler, M., Hutzler, F., & Wimmer, H. (2002). Dyslexia: verbal impairments in the absence of magnocellular impairments. *Neuroreport*, *13*(5), 617–20. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11973457

Kujala, T., Karma, K., Ceponiene, R., Belitz, S., Turkkila, P., Tervaniemi, M., & Näätänen, R. (2001). Plastic neural changes and reading improvement caused by audiovisual training in reading-impaired children. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(18), 10509–14. doi:10.1073/pnas.181589198

Kushnerenko, E., Teinonen, T., Volein, A., & Csibra, G. (2008). Electrophysiological evidence of illusory audiovisual speech percept in human infants. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(32), 11442–5. doi:10.1073/pnas.0804275105

Laasonen, M., Service, E., & Virsu, V. (2001). Temporal order and processing acuity of visual, auditory, and tactile perception in developmentally dyslexic young adults. *Cognitive, Affective, & Behavioral Neuroscience*, *1*(4), 394–410. doi:10.3758/CABN.1.4.394

Laasonen, M., Service, E., & Virsu, V. (2002). Crossmodal temporal order and processing acuity in developmentally dyslexic young adults. *Brain and Language*, *80*(3), 340–54. doi:10.1006/brln.2001.2593

Lewald, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Brain Research. Cognitive Brain Research*, *16*(3), 468–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12706226

Love, S. a, Petrini, K., Cheng, A., & Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PloS One*, *8*(1), e54798. doi:10.1371/journal.pone.0054798

Luca, M. Di, Machulla, T., & Ernst, M. O. (2009). Recalibration of multisensory simultaneity : Cross-modal transfer coincides with a change in perceptual latency. *Journal of Vision*, *9*, 1–16. doi:10.1167/9.12.7.Introduction

Lux, S., Marshall, J. C., Ritzl, A., Zilles, K., & Fink, G. R. (2003). Neural mechanisms associated with attention to temporal synchrony versus spatial orientation: an fMRI study. *NeuroImage*, *20*, S58–S65. doi:10.1016/j.neuroimage.2003.09.009

Lyon, G., Shaywitz, S., & Shaywitz, B. (2003). Defining dyslexia, comorbidity, teachers' knowledge of language and reading. *Annals of Dyslexia*. Retrieved from http://www.hku.hk/linguist/cou/adv/ling6022/articles/lyon2003.pdf

Lyytinen, H., Ronimus, M., Alanko, A., Poikkeus, A.-M., & Taanila, M. (2007). Early identification of dyslexia and the use of computer game-based practice to support reading acquisition. *Nordic Psychology*, *59*(2), 109–126. doi:10.1027/1901-2276.59.2.109

Magnan, A., Ecalle, J., Veuillet, E., & Collet, L. (2004). The effects of an audio-visual training program in dyslexic children. *Dyslexia (Chichester, England)*, *10*(2), 131–40. doi:10.1002/dys.270

Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Petersen, a. (1996). On the bases of two subtypes of developmental [corrected] dyslexia. *Cognition*, *58*(2), 157–95. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8820386

Martin, B., Giersch, A., Huron, C., & van Wassenhove, V. (2012). Temporal event structure and timing in schizophrenia: Preserved binding in a longer “now.” *Neuropsychologia*, *51*(2), 358–371. doi:10.1016/j.neuropsychologia.2012.07.002

Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. *Annual Review of Neuroscience*, *27*, 307–40. doi:10.1146/annurev.neuro.27.070203.144247

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. doi:10.1038/264746a0

Meck, W. H. (2005). Neuropsychology of timing and time perception. *Brain and Cognition*, *58*(1), 1–8. doi:10.1016/j.bandc.2004.09.004

Meng, X., Sai, X., Wang, C., Wang, J., Sha, S., & Zhou, X. (2005). Auditory and speech processing and reading development in Chinese school children: behavioural and ERP evidence. *Dyslexia*, *11*(4), 292–310. doi:10.1002/dys.309

Meredith, M., & Nemitz, J. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of*, *7*(October), 3215–3229. Retrieved from http://www.jneurosci.org/content/7/10/3215.short

Miyazaki, M., Yamamoto, S., Uchida, S., & Kitazawa, S. (2006). Bayesian calibration of simultaneity in tactile temporal order judgment. *Nature Neuroscience*, *9*(7), 875–7. doi:10.1038/nn1712

Mohammed, T., Campbell, R., Macsweeney, M., Barry, F., & Coleman, M. (2006). Speechreading and its association with reading among deaf, hearing and dyslexic individuals. *Clinical Linguistics & Phonetics*, *20*(7-8), 621–30. doi:10.1080/02699200500266745

Mongillo, E. a, Irwin, J. R., Whalen, D. H., Klaiman, C., Carter, A. S., & Schultz, R. T. (2008). Audiovisual processing in children with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(7), 1349–58. doi:10.1007/s10803-007-0521-y

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Brain Research. Cognitive Brain Research*, *17*(1), 154–63. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12763201

Moser, D., Baker, J. M., Sanchez, C. E., Rorden, C., & Fridriksson, J. (2009). Temporal order processing of syllables in the left parietal lobe. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(40), 12568–73. doi:10.1523/JNEUROSCI.5934-08.2009

Munhall, K. G., Gribble, P., Sacco, L., & Ward, M. (1996). Temporal constraints on the McGurk effect. *Perception & Psychophysics*, *58*(3), 351–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8935896

Nath, A. R., & Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *NeuroImage*, *59*(1), 781–7. doi:10.1016/j.neuroimage.2011.07.024

Navarra, J., Alsius, A., Velasco, I., Soto-Faraco, S., & Spence, C. (2010). Perception of audiovisual speech synchrony for native and non-native language. *Brain Research*, *1323*, 84–93. doi:10.1016/j.brainres.2010.01.059

Navarra, J., Hartcher-O'Brien, J., Piazza, E., & Spence, C. (2009). Adaptation to audiovisual asynchrony modulates the speeded detection of sound. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(23), 9169–73. doi:10.1073/pnas.0810486106

Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Brain Research. Cognitive Brain Research*, *25*(2), 499–507. doi:10.1016/j.cogbrainres.2005.07.009

Nenadic, I., Gaser, C., Volz, H.-P., Rammsayer, T., Häger, F., & Sauer, H. (2003). Processing of temporal information and the basal ganglia: new evidence from fMRI. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *148*(2), 238–46. doi:10.1007/s00221-002-1188-4

Nicholls, M. (1994). Hemispheric asymmetries for temporal resolution: A signal detection analysis of threshold and bias. *The Quarterly Journal of Experimental Psychology*, (January 2013), 37–41. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/14640749408401113

Nicholls, M. E. R., Gora, J., & Stough, C. K. K. (2002). Hemispheric asymmetries for visual and auditory temporal processing: an evoked potential study.

*International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *44*(1), 37–55. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11852156

Nicol, J. R., & Shore, D. I. (2007). Perceptual grouping impairs temporal resolution. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *183*(2), 141–8. doi:10.1007/s00221-007-1034-9

Nicolson, R. I., & Fawcett, A. J. (1993). Children with dyslexia automatize temporal skills more slowly. *ANNALS-NEW YORK ACADEMY OF SCIENCES*, *682*, 390–390. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.45.4840&amp;rep=rep1&amp;type=pdf

Nittrouer, S. (1999). Do temporal processing deficits cause phonological processing problems? *Journal of Speech, Language and Hearing Research*, *42*(4), 925–942. Retrieved from http://jslhr.asha.org/cgi/content/abstract/42/4/925

Noesselt, T., Bergmann, D., Heinze, H.-J., Münte, T., & Spence, C. (2012). Coding of multisensory temporal patterns in human superior temporal sulcus. *Frontiers in Integrative Neuroscience*, *6*(August), 64. doi:10.3389/fnint.2012.00064

Ojanen, V., Möttönen, R., Pekkola, J., Jääskeläinen, I. P., Joensuu, R., Autti, T., & Sams, M. (2005). Processing of audiovisual speech in Broca's area. *NeuroImage*, *25*(2), 333–8. doi:10.1016/j.neuroimage.2004.12.001

Okubo, M., & Nicholls, M. E. R. (2008). Hemispheric asymmetries for temporal information processing: transient detection versus sustained monitoring. *Brain and Cognition*, *66*(2), 168–75. doi:10.1016/j.bandc.2007.07.002

Olulade, O. a, Napoliello, E. M., & Eden, G. F. (2013). Abnormal visual motion processing is not a cause of dyslexia. *Neuron*, *79*(1), 180–90. doi:10.1016/j.neuron.2013.05.002

Pekkola, J., Laasonen, M., Ojanen, V., Autti, T., Jääskeläinen, I. P., Kujala, T., & Sams, M. (2006). Perception of matching and conflicting audiovisual speech in dyslexic and fluent readers: an fMRI study at 3 T. *NeuroImage*, *29*(3), 797–807. doi:10.1016/j.neuroimage.2005.09.069

Petrini, K., Dahl, S., Rocchesso, D., Waadeland, C. H., Avanzini, F., Puce, A., & Pollick, F. E. (2009). Multisensory integration of drumming actions: musical expertise affects perceived audiovisual asynchrony. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *198*(2-3), 339–52. doi:10.1007/s00221-009-1817-2

Pöppel, E. (1988). *Mindworks: Time and conscious experience* (p. 211). Harcourt Brace Jovanovich.

Powers, A. R., Hillock, A. R., & Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(39), 12265–74. doi:10.1523/JNEUROSCI.3501-09.2009

Ramirez, J., & Mann, V. (2005). Using auditory-visual speech to probe the basis of noise-impaired consonant–vowel perception in dyslexia and auditory neuropathy. *The Journal of the Acoustical Society of America*, *118*(2), 1122. doi:10.1121/1.1940509

Rammsayer, T. H., & Brandler, S. (2007). Performance on temporal information processing as an index of general intelligence. *Intelligence*, *35*(2), 123–139. doi:10.1016/j.intell.2006.04.007

Ramus, F. (2001). Outstanding questions about phonological processing in dyslexia. *Dyslexia (Chichester, England)*, *7*(4), 197–216. doi:10.1002/dys.205

Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology*, *13*(2), 212–218. doi:10.1016/S0959-4388(03)00035-7

Rapcsak, S., Beeson, P., Henry, M., & Leyden, A. (2009). Phonological dyslexia and dysgraphia: Cognitive mechanisms and neural substrates. *Cortex*, *45*(5), 575–591. doi:10.1016/j.cortex.2008.04.006.Phonological

Raschle, N. M., Stering, P. L., Meissner, S. N., & Gaab, N. (2013). Altered Neuronal Response During Rapid Auditory Processing and Its Relation to Phonological Processing in Prereading Children at Familial Risk for Dyslexia. *Cerebral Cortex (New York, N.Y. : 1991)*, (Ramus 2003). doi:10.1093/cercor/bht104

Reed, M. A. (1989). Speech perception and the discrimination of brief auditory cues in reading disabled children. *Journal of Experimental Child Psychology*, *48*(2), 270–292. Retrieved from http://www.sciencedirect.com/science/article/pii/0022096589900064

Rey, V., De Martino, S., Espesser, R., & Habib, M. (2002). Temporal processing and phonological impairment in dyslexia: effect of phoneme lengthening on order judgment of two consonants. *Brain and Language*, *80*(3), 576–91. doi:10.1006/brln.2001.2618

Roach, N. W., Heron, J., Whitaker, D., & McGraw, P. V. (2011). Asynchrony adaptation reveals neural population code for audio-visual timing. *Proceedings. Biological Sciences / The Royal Society*, *278*(1710), 1314–22. doi:10.1098/rspb.2010.1737

Roseboom, W., & Arnold, D. H. (2011). Twice Upon a Time: Multiple Concurrent Temporal Recalibrations of Audiovisual Speech. *Psychological Science*, (June). doi:10.1177/0956797611413293

Roseboom, W., Nishida, S., & Arnold, D. H. (2009). The sliding window of audio–visual simultaneity. *Journal of Vision*, *9*(12), 1–8. doi:10.1167/9.12.4.Introduction

Roseboom, W., Nishida, S., Fujisaki, W., & Arnold, D. H. (2011). Audio-visual speech timing sensitivity is enhanced in cluttered conditions. *PloS One*, *6*(4), e18309. doi:10.1371/journal.pone.0018309

Rosen, S. (2003). Auditory processing in dyslexia and specific language impairment: is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics*, *31*(3-4), 509–527. doi:10.1016/S0095-4470(03)00046-9

Rosen, S., & Manganari, E. (2001). Is there a relationship between speech and nonspeech auditory processing in children with dyslexia? *Journal of Speech, Language and Hearing …*, (August). Retrieved from http://jslhr.asha.org/cgi/content/abstract/44/4/720

Rosenblum, L. D., Schmuckler, M. a, & Johnson, J. a. (1997). The McGurk effect in infants. *Perception & Psychophysics*, *59*(3), 347–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9136265

Rubia, K., & Smith, A. (2004). The neural correlates of cognitive time management: a review. *Acta Neurobiologiae Experimentalis*, *64*(3), 329–40. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15283476

Saint-Amour, D., De Sanctis, P., Molholm, S., Ritter, W., & Foxe, J. J. (2007). Seeing voices: High-density electrical mapping and source-analysis of the multisensory mismatch negativity evoked during the McGurk illusion. *Neuropsychologia*, *45*(3), 587–97. doi:10.1016/j.neuropsychologia.2006.03.036

Sale, A., Berardi, N., & Maffei, L. (2012). Environmental Influences on Visual Cortex Development and Plasticity. In S. Molotchnikoff & J. Rouat (Eds.), *Visual Cortex – Current Status and Perspectives* (pp. 295–322). InTech.

Savage, R. S., Frederickson, N., Goodwin, R., Patni, U., Smith, N., & Tuersley, L. (2005). Relationships Among Rapid Digit Naming, Phonological Processing, Motor Automaticity, and Speech Perception in Poor, Average, and Good Readers and Spellers. *Journal of Learning Disabilities*, *38*(1), 12–28. doi:10.1177/00222194050380010201

Schneider, K. a., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology*, *47*(4), 333–366. doi:10.1016/S0010-0285(03)00035-5

Schulte-körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1998). Auditory processing and dyslexia : evidence for a specific speech, *9*(2), 337–340.

Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1999). The role of phonological awareness, speech perception, and auditory temporal processing for dyslexia. *European Child & Adolescent Psychiatry*, *8 Suppl 3*, 28–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10638366

Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (2001). Speech perception deficit in dyslexic adults as measured by mismatch negativity (MMN). *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *40*(1), 77–87. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11166109

Sekiyama, K. (2003). Auditory-visual speech perception examined by fMRI and PET. *Neuroscience Research*, *47*(3), 277–287. doi:10.1016/S0168-0102(03)00214-1

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*(6614), 308. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9002513

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *NeuroImage*, *25*(1), 76–89. doi:10.1016/j.neuroimage.2004.11.006

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex (New York, N.Y. : 1991)*, *17*(10), 2387–99. doi:10.1093/cercor/bhl147

Slutsky, D., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*, *12*(1), 7–10. Retrieved from http://journals.lww.com/neuroreport/Abstract/2001/01220/Temporal_and_spatial_dependency_of_the.9.aspx

Soto-Faraco, S., & Alsius, A. (2007). Conscious access to the unisensory components of a cross-modal illusion. *Neuroreport*, *18*(4), 347–50. doi:10.1097/WNR.0b013e32801776f9

Soto-Faraco, S., & Alsius, A. (2009). Deconstructing the McGurk–MacDonald illusion. *Journal of Experimental Psychology: …*, 1–8. doi:10.1037/a0013483

Spence, C. (2007). Audiovisual multisensory integration. *Acoustical Science and Technology*, *28*(2), 61–70. doi:10.1250/ast.28.61

Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: when two locations are better than one.

*Perception & Psychophysics*, *65*(2), 318–28. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12713247

Spence, C., & Parise, C. (2010). Prior-entry: a review. *Consciousness and Cognition*, *19*(1), 364–79. doi:10.1016/j.concog.2009.12.001

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*(4), 799–832. doi:10.1037//0096-3445.130.4.799

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

Stanovich, K. E. (1988). Explaining the Differences Between the Dyslexic and the Garden-Variety Poor Reader: The Phonological-Core Variable-Difference Model. *Journal of Learning Disabilities*, *21*(10), 590–604. doi:10.1177/002221948802101003

Stein, J. (2001). The magnocellular theory of developmental dyslexia. *Dyslexia (Chichester, England)*, *7*(1), 12–36. doi:10.1002/dys.186

Stein, J., & Talcott, J. (1999). Impaired Neuronal Timing in Developmental Dyslexia—The Magnocellular Hypothesis. *Dyslexia*, *77*, 59–77. Retrieved from http://www.physiol.ox.ac.uk/~jfs/pdf/dyslexiaj.pdf

Stein, J., & Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in Neurosciences*, *20*(4), 147–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9106353

Stevenson, R. a, & James, T. W. (2009). Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *NeuroImage*, *44*(3), 1210–23. doi:10.1016/j.neuroimage.2008.09.034

Stevenson, R. a, VanDerKlok, R. M., Pisoni, D. B., & James, T. W. (2011). Discrete neural substrates underlie complementary audiovisual speech integration processes. *NeuroImage*, *55*(3), 1339–45. doi:10.1016/j.neuroimage.2010.12.063

Stevenson, R. A., Altieri, N. A., Kim, S., Pisoni, D. B., & James, T. W. (2010). Neural processing of asynchronous audiovisual speech perception. *NeuroImage*, *49*(4), 3308–3318. doi:10.1016/j.neuroimage.2009.12.001

Stone, J. V, Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., … Porter, N. R. (2001). When is now? Perception of simultaneity. *Proceedings. Biological Sciences / The Royal Society*, *268*(1462), 31–8. doi:10.1098/rspb.2000.1326

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*, *421*(6926), 911. doi:10.1038/421911a

Summerfield, Q. (1992). Lipreading and audio-visual speech perception. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *335*(1273), 71–8. doi:10.1098/rstb.1992.0009

Szycik, G. R., Stadler, J., Tempelmann, C., & Münte, T. F. (2012). Examining the McGurk illusion using high-field 7 Tesla functional MRI. *Frontiers in Human Neuroscience*, *6*(April), 95. doi:10.3389/fnhum.2012.00095

Talcott, J. B., Hansen, P. C., Assoku, E. L., & Stein, J. F. (2000). Visual motion sensitivity in dyslexia: evidence for temporal and energy integration deficits. *Neuropsychologia*, *38*(7), 935–43. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10775704

Tallal, P. (1980a). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*(2), 182–198. doi:10.1016/0093-934X(80)90139-X

Tallal, P. (1980b). Language and reading: Some perceptual prerequisites. *Bulletin of the Orton Society*, *30*, 170–178. Retrieved from http://link.springer.com/article/10.1007/BF02653716

Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews. Neuroscience*, *5*(9), 721–8. doi:10.1038/nrn1499

Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. *Annals of the New York Academy of Sciences*, *682*(1), 27–47. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.1993.tb22957.x/abstract

Taylor, N., Isaac, C., & Milne, E. (2010). A comparison of the development of audiovisual integration in children with autism spectrum disorders and typically developing children. *Journal of Autism and Developmental Disorders*, *40*(11), 1403–11. doi:10.1007/s10803-010-1000-4

Teinonen, T., Aslin, R. N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition*, *108*(3), 850–5. doi:10.1016/j.cognition.2008.05.009

Thompson, P. M., Cannon, T. D., Narr, K. L., van Erp, T., Poutanen, V. P., Huttunen, M., … Toga, a W. (2001). Genetic influences on brain structure. *Nature Neuroscience*, *4*(12), 1253–8. doi:10.1038/nn758

Tsermentseli, S., O'Brien, J. M., & Spencer, J. V. (2008). Comparison of form and motion coherence processing in autistic spectrum disorders and dyslexia. *Journal of Autism and Developmental Disorders*, *38*(7), 1201–10. doi:10.1007/s10803-007-0500-3

Valente, A. a, Miguel, E. C., Castro, C. C., Amaro, E., Duran, F. L. S., Buchpiguel, C. a, … Busatto, G. F. (2005). Regional gray matter abnormalities in obsessive-compulsive disorder: a voxel-based morphometry study. *Biological Psychiatry*, *58*(6), 479–87. doi:10.1016/j.biopsych.2005.04.021

Van Atteveldt, N. M., Formisano, E., Blomert, L., & Goebel, R. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cerebral Cortex (New York, N.Y. : 1991)*, *17*(4), 962–74. doi:10.1093/cercor/bhl007

Van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & Van De Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception & Psychophysics*, *70*(6), 955–968. doi:10.3758/PP.70.6.955

Van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & Van De Par, S. (2010). Temporal order judgment criteria are affected by synchrony judgment sensitivity. *Attention, Perception, & Psychophysics*, *72*(8), 2227. doi:10.3758/APP

Van Gaal, S., Scholte, H. S., Lamme, V. a F., Fahrenfort, J. J., & Ridderinkhof, K. R. (2011). Pre-SMA graymatter density predicts individual differences in action selection in the face of conscious and unconscious response conflict. *Journal of Cognitive Neuroscience*, *23*(2), 382–90. doi:10.1162/jocn.2010.21444

Van Ingelghem, M., van Wieringen, A., Wouters, J., Vandenbussche, E., Onghena, P., & Ghesquière, P. (2001). Psychophysical evidence for a general temporal processing deficit in children with dyslexia. *Neuroreport*, *12*(16), 3603–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11733720

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45*(3), 598–607. doi:10.1016/j.neuropsychologia.2006.01.001

Vandermosten, M., Boets, B., Luts, H., Poelmans, H., Golestani, N., Wouters, J., & Ghesquière, P. (2010). Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(23), 10389–94. doi:10.1073/pnas.0912858107

Vandermosten, M., Boets, B., Luts, H., Poelmans, H., Wouters, J., & Ghesquière, P. (2011). Impairments in speech and nonspeech sound categorization in children with dyslexia are driven by temporal processing difficulties. *Research in Developmental Disabilities*, *32*(2), 593–603. doi:10.1016/j.ridd.2010.12.015

Vatakis, A., Ghazanfar, A. A., & Spence, C. (2008). Facilitation of multisensory integration by the "unity effect" reveals that speech is special. *Journal of Vision*, *8*(9), 14.1–11. doi:10.1167/8.9.14

Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Temporal recalibration during asynchronous audiovisual speech perception. *Experimental Brain Research.*, *181*(1), 173–81. doi:10.1007/s00221-007-0918-z

Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2008). Audiovisual temporal adaptation of speech: temporal order versus simultaneity judgments. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *185*(3), 521–9. doi:10.1007/s00221-007-1168-9

Vatakis, A., & Spence, C. (2006). Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. *Neuroscience Letters*, *393*(1), 40–4. doi:10.1016/j.neulet.2005.09.032

Vatakis, A., & Spence, C. (2007). Crossmodal binding: evaluating the "unity assumption" using audiovisual speech stimuli. *Perception & Psychophysics*, *69*(5), 744–56. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17929697

Vatakis, A., & Spence, C. (2008a). Evaluating the influence of the "unity assumption" on the temporal perception of realistic audiovisual stimuli. *Acta Psychologica*, *127*(1), 12–23. doi:10.1016/j.actpsy.2006.12.002

Vatakis, A., & Spence, C. (2008b). Investigating the effects of inversion on configural processing with an audiovisual temporal-order judgment task. *Perception*, *37*(1), 143–160. doi:10.1068/p5648

Veuillet, E., Magnan, A., Ecalle, J., Thai-Van, H., & Collet, L. (2007). Auditory processing disorder in children with reading disabilities: effect of audiovisual training. *Brain : A Journal of Neurology*, *130*(Pt 11), 2915–28. doi:10.1093/brain/awm235

Virsu, V., Lahti-Nuuttila, P., & Laasonen, M. (2003). Crossmodal temporal processing acuity impairment aggravates with age in developmental dyslexia. *Neuroscience Letters*, *336*, 151–154. doi:10.1016/S03

Vroomen, J., & Keetels, M. (2010a). Perception of intersensory synchrony: a tutorial review. *Attention Perception Psychophysics*, *72*(4), 871–884. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20436185

Vroomen, J., & Keetels, M. (2010b). Perception of intersensory synchrony: a tutorial review. *Attention Perception Psychophysics*, *72*(4), 871–884. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20436185

Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Brain Research. Cognitive Brain Research*, *22*(1), 32–5. doi:10.1016/j.cogbrainres.2004.07.003

Vroomen, J., & Stekelenburg, J. J. (2011). Perception of intersensory synchrony in audiovisual speech: not that special. *Cognition*, *118*(1), 75–83. doi:10.1016/j.cognition.2010.10.002

Weiss, K., & Scharlau, I. (2011). Simultaneity and temporal order perception: Different sides of the same coin? Evidence from a visual prior-entry study. *Quarterly Journal of Experimental Psychology (2006)*, *64*(2), 394–416. doi:10.1080/17470218.2010.495783

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–67. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7003641

Westlye, L. T., Grydeland, H., Walhovd, K. B., & Fjell, A. M. (2011). Associations between regional cortical thickness and attentional networks as measured by the attention network test. *Cerebral Cortex (New York, N.Y. : 1991)*, *21*(2), 345–56. doi:10.1093/cercor/bhq101

Widmann, A., Schröger, E., Tervaniemi, M., Pakarinen, S., & Kujala, T. (2012). Mapping symbols to sounds: electrophysiological correlates of the impaired reading process in dyslexia. *Frontiers in Psychology*, *3*(March), 60. doi:10.3389/fpsyg.2012.00060

Wiener, M., Matell, M. S., & Coslett, H. B. (2011). Multiple mechanisms for temporal processing. *Frontiers in Integrative Neuroscience*, *5*(July), 31. doi:10.3389/fnint.2011.00031

Wittmann, M., & Burtscher, A. (2004). Effects of brain-lesion size and location on temporal-order judgment. *Neuroreport*, *15*(15), 1–5. Retrieved from http://journals.lww.com/neuroreport/Abstract/2004/10250/Effects_of_brain_lesion_size_and_location_on.20.aspx

Witton, C., Talcott, J. B., Hansen, P. C., Richardson, a J., Griffiths, T. D., Rees, A., … Green, G. G. (1998). Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers. *Current Biology : CB*, *8*(14), 791–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9663387

Wright, T. M., Pelphrey, K. a, Allison, T., McKeown, M. J., & McCarthy, G. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cerebral Cortex (New York, N.Y. : 1991)*, *13*(10), 1034–43. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12967920

Yamamoto, S., Miyazaki, M., Iwano, T., & Kitazawa, S. (2012). Bayesian calibration of simultaneity in audiovisual temporal order judgments. *PloS One*, *7*(7), e40379. doi:10.1371/journal.pone.0040379

Yarrow, K., Jahn, N., Durant, S., & Arnold, D. H. (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Consciousness and Cognition*, *20*(4), 1518–31. doi:10.1016/j.concog.2011.07.003

Zampini, M. (2003). Multisensory temporal order judgments: the role of hemispheric redundancy. *International Journal of Psychophysiology*, *50*(1-2), 165–180. doi:10.1016/S0167-8760(03)00132-6

Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neuroscience Letters*, *381*(3), 217–22. doi:10.1016/j.neulet.2005.01.085

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, *12*(5), 732–45. doi:10.1111/j.1467-7687.2009.00817.x