MECHANISMS OF AUDIOVISUAL INTEGRATION IN THE YOUNG AND HEALTHY AGEING BRAIN

by SAMUEL ASHLEY JONES

A thesis submitted to the University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology
College of Life and Environmental Sciences
University of Birmingham
September 2018

ABSTRACT

Throughout our lifespan we are continually exposed to a barrage of multisensory signals. It is only by successfully resolving these signals into a veridical percept that we are able to interact effectively with the world. This crucial process of multisensory integration has previously been shown to change as we age, but the mechanisms underlying these age differences are not well understood. The work described in this thesis applied behavioural testing, psychophysics, computational modelling, and functional MRI to improve our understanding of the relationship between healthy brain ageing and audiovisual integration. Following a review of the existing literature (Chapter 1) and a description of methodology (Chapter 2), I describe a study that utilised Bayesian modelling to describe age differences in responses to a spatial ventriloquist paradigm (Chapter 3). This study revealed that older adults ultimately respond in a similar way to younger adults when presented with audiovisual spatial cues, but reaction time analyses indicated the possibility of different underlying mechanisms. The research outlined in Chapter 4 thus used multivariate analysis of functional MRI data to investigate the cortical networks underpinning audiovisual spatial integration in older adults, finding that these remain largely unchanged throughout the adult lifespan. In Chapter 5 I describe a study that instead applied functional MRI to the investigation of audiovisual speech comprehension and its interactions with age, asynchrony, and background noise. We show that older adults' speech comprehension is more impaired by the presence of background noise and stimulus asynchrony, but that their perception of the asynchrony itself is intact. We demonstrate neural effects that account for these age differences. Finally, in Chapter 6 I review these various findings and explore their implications in relation to previous and future research.

ACKNOWLEDGEMENTS

There are a great many people without whom this thesis would not exist, and I will aim to thank as many of them as possible here. First my supervisor, Uta Noppeney, whose relentless dedication to research has pushed me to achieve in areas that I would have never previously considered, and our various collaborators who improved the standard of this research significantly. Further thanks to the many administrative staff, academics, and students in Birmingham who improved my time here tremendously.

Next, I would like to thank all past and present members of the CCN lab. You are an amazing, talented, and unique group of people who have helped me in more ways than I can express. Specific thanks to Arianna Zuanazzi, who started when I did and showed off by finishing a day early. I'm glad you were around and will probably forgive you eventually. To Ambra Ferrari, who started far later yet still knows more than me (though she'll never believe it). To Tom White, who somehow turned overnight MRI scanning into a recreational activity. To Máté Aller, who provided an outlet for deeply nerdy tendencies that others don't know how to handle. To Remi Gau who, despite an extended absence in sunny paradise, remains a constant source of both cynical fun and supportive advice. And, of course, to David Meijer, who is a complete and invaluable friend with a distressingly relaxed attitude towards traffic signs. Agoston, Giulio, Steffen, Patrycja, Johanna, Michael, Alex, and Alex: thank you all.

I thank my mum, Julie Jones, and my sister Becky Jones, for accepting my strange decision to run away and become a scientist and for frequently checking that it hasn't yet killed me, and I thank the rest of my family for doing their very best to be interested in too many brain-related anecdotes. Dad, Charlie, Martha, Nan, and Grandad: thank you.

Thank you to old friends, Colum Taylor, Chris Thomas, Claire Woolgrove, Emily Keen, and so many others, who have had the job of keeping me in touch with reality for far too long.

I must also thank my endlessly dedicated and tolerant elderly participants, many of whom would be deeply offended if they heard me call them that. They have, without exception, been a pleasure to work with.

Finally, Ella Kirsty Hawkins. You bring joy, optimism, wit, intelligence, unwavering support, nourishment, and doggos to my life. Without you, more than anyone, this would not have happened. Thank you.

Now on to the science.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS

A1	Primary auditory cortex
AES	Anterior ectosylvian sulcus
ANOVA	Analysis of variance
BADS	Bayesian adaptive direct search
BCI	Bayesian causal inference
BOLD	Blood oxygenation level dependent
CMB	Crossmodal bias
CO2	Carbon dioxide
CSF	Cerebrospinal fluid
dB	Decibels
DCM	Dynamic causal modelling
ECoG	Electrocorticography
EEG	Electroencephalography
EPI	Echo planar imaging
FEF	Frontal eye field
fMRI	Functional magnetic resonance imaging
FWE	Familywise error
FWHM	Full width at half maximum
GLM	General linear model
HAROLD	Hemispheric asymmetry reduction in older adults
HL	Hearing level
HRF	Haemodynamic response function
IFG	Inferior frontal gyrus
IPS	Intraparietal sulcus
JND	Just noticeable difference
MEG	Magnetoencephalography
MLE	Maximum likelihood estimation
MNI	Montreal Neurological Institute
mSTC	Multisensory superior temporal cortex
MVB	Multivariate Bayes
MVPA	Multivariate pattern analysis
PASA	Posterior to anterior shift in ageing
PFC	Prefrontal cortex
PSE	Point of subjective equality
PT	Planum temporale
RMSE	Root mean squared error
ROI	Region of interest
RSFA	Resting state fluctuation amplitude
RT	Response time
SEM	Standard error of the mean
SNR	Signal to noise ratio
SPL	Sound pressure level
SPM	Statistical parametric mapping
STAC	Scaffolding theory of ageing and cognition
STC	Superior temporal cortex
STG	Superior temporal gyrus
STS	Superior temporal sulcus
SVM	Support vector machine
SVR	Support vector regression
TA	Acquisition time
TR	Repetition time
V1-V3	Visual areas 1 to 3
VIP	Ventral intraparietal area

CHAPTER 1: GENERAL INTRODUCTION

We exist in a world that is infinitely variable. Any object within it can be characterised according to a great number of properties: size, mass, position, velocity, transparency, temperature, elasticity, viscosity. It is only by estimating these properties that an organism can respond effectively to its environment. To this end, humans—like most other organisms—have evolved an array of senses with a variety of capabilities. Crucially, these senses do not operate in isolation, but combine with and complement each other. For example, our sense of hearing is generally sufficient for understanding another person's speech in a quiet room, but quickly becomes less useful as the environment becomes more challenging, such as in a noisy bar. In such cases, the addition of visual information (in the form of lip-reading) can make the difference between understanding your conversation partner and not. This process of combining multiple senses is referred to broadly as multisensory integration, and it is central to our ability to interact with the world around us.

As with most physical or mental capabilities, our central nervous system's ability to perform this process does not remain static across the lifespan. Multisensory integration in its various forms develops gradually throughout childhood (Burr & Gori, 2012). While young infants are aware of amodal properties of objects from early infancy (Streri & Gentaz, 2003), for instance, and multisensory facilitation of reflexive movements is apparent from around eight months (Neil, Chee-Ruiter, Scheier, Lewkowicz, & Shimojo, 2006), perception of multisensory speech illusions does not begin to match that of adults until around 10 years of age (Tremblay et al., 2007). The focus of this thesis, however, is the changes to multisensory integration processes that may occur towards the later years of life. In this introductory chapter I review our current understanding of multisensory integration as a neural, perceptual, and cognitive process, before considering the ways in which healthy ageing can change the human brain and the resulting cognitive and behavioural outcomes. These are each vast areas of research that it would be impossible for me to review exhaustively; I instead give an overview of some key theories and findings, focusing particularly on those that are most relevant to the empirical work described in later chapters. I conclude this chapter by discussing previous research that has investigated the effects of healthy ageing on multisensory integration.

Multisensory integration

Our ability to understand, and therefore react to, our environment is greatly enhanced when we combine multiple sensory signals. The effective combination of these signals is, however, a computationally complex problem for the sensory system to solve. First, it must decide which of the thousands of simultaneous inputs belong together—that is, have the same source. This is known as causal inference, and it is challenging because the brain has no direct access to the true state of the world; it can only make inferences based upon the interpretation of noisy sensory data. The relative contribution of those signals that have been determined to have a common source must then be integrated in a way that maximises the usefulness of the combined estimate. Despite this complexity, behavioural (Alais & Burr, 2004; Ernst & Banks, 2002; Jacobs, 1999; Koerding et al., 2007; Rohe & Noppeney, 2015a) and neural (Fetsch, DeAngelis, & Angelaki, 2013; Fetsch, Pouget, DeAngelis, & Angelaki, 2012; Ma, Beck, Latham, & Pouget, 2006; Morgan, DeAngelis, & Angelaki, 2008; Rohe & Noppeney, 2015b, 2016, 2018) modelling has demonstrated that young, healthy adults generally perform in a way that can be closely approximated by mathematically optimal Maximum Likelihood Estimation (MLE) and Bayesian Causal Inference (BCI) models, suggesting that the brain is highly optimised for this process. I discuss these models in detail in Chapter 2. In this section I will outline our current understanding of the neural mechanisms that drive these operations and review their behavioural outcomes, before considering how other cognitive processes might interact with multisensory integration.

Integration in the single cell

Some of the earliest work to investigate the perceptual process of multisensory integration was conducted by Stein and colleagues and focused on single-cell recordings in superior colliculus of non-human mammals (Meredith & Stein, 1983; Wallace, Wilkinson, & Stein, 1996). The researchers observed that the response of some cells to multisensory stimuli was superadditive, i.e. greater than the sum of their responses to unisensory stimuli. Importantly, the effect was only apparent when the auditory and visual stimuli were temporally coincident and appeared within the same area of space. Stein and Stanford (2008) argue that this multisensory enhancement of a neuron's response reflects an increase in the physiological salience of an event and serves to facilitate faster motor responses. The researchers also noted depression in the responses of some neurons when stimuli from multiple sensory modalities were presented simultaneously but in spatially disparate locations, perhaps reflecting a mechanism for segregation of unrelated stimuli (Noppeney, Jones, Rohe, & Ferrari, 2018).

The same cell recording techniques have revealed the existence of similarly multisensory cells in the cortex (Wallace, Meredith, & Stein, 1993). In the cat, these are primarily in anterior ectosylvian sulcus (AES), an area that also contains many unisensory cells that project to superior colliculus (Stein & Wallace, 1996). This area does not, however, have a clear homologue in primates, so its utility for making inferences about human processing is limited. Instead, cortical single-cell recording of responses to multisensory stimuli in primates focuses mainly on posterior parietal cortex, superior temporal sulcus (STS), and prefrontal cortex (PFC). Multisensory interactions in these areas are often more nuanced than those observed in superior colliculus (or in cat AES). Avillac, Hamed, and Duhamel (2007), for example, demonstrate that most neurons in ventral intraparietal area (VIP) of the macaque monkey are in some way multisensory, but respond in a variety of nonlinear additive, subadditive, and superadditive ways to temporally and spatially coincident visuotactile stimuli. The importance of STS for speech integration in humans, discussed in detail later, is also already hinted at in macaque single-unit recordings: Barraclough, Xiao, Baker, Oram, and Perrett (2005) found that monkey vocalisations presented together with congruent videos of human face movement elicited greater multisensory enhancement in STS neurons than those that were matched with incongruent videos. An area of macaque PFC has also been shown to be concerned with the integration of audiovisual communication signals (Sugihara, Diltz, Averback, & Romanski, 2006).

Integration across the cortex

Investigations of multisensory integration in the cortex are not confined to single-cell recordings in animals. A vast and growing body of literature applies neuroimaging techniques such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), magnetoencephalography (MEG), and electrocorticography (ECoG) to the task of investigating the cortical regions and networks involved in multisensory integration in humans. The earliest work, building on extensive evidence of functional specificity in primary sensory cortices (Brewer & Barton, 2016; Grill-Spector & Malach, 2004), assumed a model wherein lower-level sensory areas processed their respective unisensory signals before feeding to higher-level areas for integration. Multisensory interactions have indeed been demonstrated in a number of higher-order areas in both humans and animals, including superior temporal sulcus (Barraclough, et al., 2005; Beauchamp, Lee, Argall, & Martin, 2004; Calvert, Campbell, & Brammer, 2000; Miller & D’Esposito, 2005; Stevenson & James, 2009), intraparietal sulcus (Avillac, Ben Hamed, & Duhamel, 2007; Rohe and Noppeney, 2015, 2016; Miller & D'Esposito, 2005), superior parietal lobule (Macaluso, Driver, & Frith, 2003; Molholm et al., 2006), and prefrontal cortex (Noppeney, Ostwald, & Werner, 2010; Sugihara, Diltz, Averbeck, & Romanski, 2006). However, it is also becoming increasingly evident that integration occurs far beyond these traditionally multisensory regions, sometimes even as early as primary sensory cortices (Foxe et al., 2002; Lehmann et al., 2006; Martuzzi et al., 2007). Ibrahim et al. (2016), for example, demonstrated that the presence of sounds sharpened the responses of orientation-selective neurons in mouse primary visual cortex. In humans, Lee and Noppeney (2014) found that asynchronous movies in which the visual stimulus preceded the auditory led to increased activation in auditory cortex, while those in which the auditory led the visual resulted in increased visual cortex activation. It is therefore reasonable to conclude that the operations on which multisensory integration depends take place across a widely-distributed network of cortical areas; Ghazanfar and Schroeder (2006) even go so far as to suggest that the entire neocortex may be considered essentially multisensory. Importantly, though, the relative contributions of the various regions are heavily dependent upon the specifics of the stimuli. Multisensory integration and multisensory perception are often discussed conceptually as general processes, but the actual operations required to solve the problem of visual-haptic shape judgement are very different from those involved in audiovisual object localisation, which is again very different from multisensory speech processing. The neural activation patterns reflect these differences.

Rohe and Noppeney's (2015, 2016, 2018) fMRI research sheds light on the cortical mechanisms underlying audiovisual spatial integration. Previous research (Avillac et al., 2007; Macaluso & Driver, 2005; Schlack et al., 2005) has shown that intraparietal sulcus is a critical region for the integration of spatial information. However, Rohe and Noppeney were able to demonstrate the existence of a cortical hierarchy wherein audiovisual spatial signals are represented in various ways, corresponding to stages of a BCI model of multisensory integration (discussed in detail in the following chapter). In a paradigm similar to that used in Chapter 4 of this thesis, participants were presented with auditory and visual stimuli at multiple degrees of spatial incongruence and instructed to report the location of either the auditory or visual stimulus. Multivariate pattern analysis (MVPA) was then used to relate activation patterns in various cortical regions to stimulus representations at each level of the BCI model. Lower-level auditory (A1) and visual (V1-V3) areas were found to separately encode the location of simultaneous auditory and visual signals (respectively). Anterior visual cortex (V3AB) and posterior intraparietal sulcus (IPS0-2) instead primarily encoded a fully fused (integrated) percept, a weighted average of the two estimated locations. Finally, anterior IPS (3-4) encoded the final spatial estimate as a weighted combination of the unisensory estimates and fused percept. Dynamic causal modelling (DCM) also revealed that prefrontal cortex (PFC) exerted top-down control over IPS, modulating its activity based on task type. The cortical network for audiovisual spatial integration therefore seems to consist of a hierarchy that includes primary sensory areas, posterior and anterior intraparietal sulcus, and prefrontal cortex. Mihalik (2017) later demonstrated that the frontal eye field (FEF) and planum temporale (PT) may also play a role. Interestingly, Rohe and Noppeney were able to decode spatial location of the non-preferred modality from primary sensory areas (i.e. auditory location from primary visual; visual location from primary auditory) at above-chance levels, demonstrating the low-level multisensory representations described earlier.

These low-level multisensory influences are also apparent for speech. Okada, Venezia, Matchin, Saberi, and Hickok (2013), for example, demonstrated that congruent visual speech stimuli increased activations relative to unisensory auditory in primary auditory cortex, an effect supported by analogous results in macaques (Ghazanfar, Maier, Hoffman, and Logothetis, 2005). Others, including Calvert et al. (1997) and Pekkola et al. (2005), have even found primary auditory activations for visual-only speech stimuli. Beyond these primary sensory regions, however, multisensory speech perception seems to rely on a mostly separate network from audiovisual localisation. Superior temporal gyrus (STG) and superior temporal sulcus (STS), sometimes collectively referred to as superior temporal cortex (STC), appears to be the main locus of activity during audiovisual speech integration (Matchin, Groulx, & Hickok, 2013; Miller & D'Esposito, 2005; Nath & Beauchamp, 2011, 2012; Stevenson & James, 2009; Stevenson, VanDerKlok, Pisoni, & James, 2010; Szycik, Jansma, & Münte, 2009; Venezia et al., 2017). Inferior frontal regions, such as Broca's area (Brodmann areas 44 & 45), may also play a role (Sams, Möttönen, & Sihvonen, 2005; Skipper, van Wassenhove, Nusbaum, & Small, 2007; Szycik et al, 2009; Watkins, Strafella, & Paus, 2003; but see Matchin et al., 2014).

Ozker, Schepers, Magnotti, Yoshor, & Beauchamp (2017) recently conducted a study that utilised ECoG to directly observe the distribution of activity along STG in response to multisensory speech stimuli. Electrical activity was measured at electrodes implanted at various points on the surface of STG in epilepsy patients. Participants were shown audiovisual movies of a person speaking single-syllable words under four conditions: normal movies with clear auditory and visual signals, movies with degraded video, movies with degraded audio, or movies with both audio and video degraded. Anterior STG was found to respond differentially to degraded versus normal auditory signals in terms of response amplitude, variability, and timing, while posterior STG responded similarly to all stimuli. A sharp boundary divided the two regions. The researchers concluded that the anterior part of STG was encoding unisensory auditory information, and the representations were thus impacted by its reliability, while posterior STG instead held an integrated audiovisual representation that compensated for degradation in the auditory. This was a novel finding using rare and valuable ECoG data, though it must be noted that there was no clear effect of visual reliability on activations in either location. The researchers' interpretation would have been strengthened by the presence of a differential response to normal versus degraded visual stimuli (in the context of degraded auditory) in posterior STG, indicating that different levels of visual information improved the estimate to varying degrees.

Behavioural responses to multisensory stimuli

These increasingly well-defined multisensory interactions that are apparent throughout the cortex serve to facilitate perception and behaviour. Multisensory stimuli allow us to respond faster (Diederich & Colonius, 2004; Forster, Cavina-Pratesi, Aglioti, & Berlucchi, 2001; Laurienti, Burdette, Maldjian, & Wallace, 2006), detect lower-intensity inputs (Frassinetti, Bolognini, & Làdavas, 2002; Gillmeister & Eimier, 2007; Noesselt, Bergmann, Hake, Heinze, & Fendrich, 2008; Vroomen & de Gelder, 2000), understand speech more clearly (Ross, Saint-Amour, Leavitt, Javitt, & Foxe, 2007; Schwartz, Berthommier, & Savariaux, 2004), recognise items more easily (Giard & Peronnet, 1999; Newell, Bülthoff, & Ernst, 2003), and locate objects more accurately (Alais & Burr, 2004; Chan, Maguinness, Lisiecka, Setti, & Newell, 2012; Hairstone, Laurienti, Mishra, Burdette, & Wallace, 2003). Some experiences are even exclusively multisensory: flavour, for example, is a percept that only arises as a combination of taste and olfaction (Dalton, Doolittle, Nagata, & Breslin, 2000; Spence, 2013a). Here I will attempt to give a short summary of the diverse research that has used psychophysical and other behavioural testing to quantify our perception of multisensory stimuli.

As discussed above, the combining of signals (between or within senses) should only occur if there is sufficient evidence that they originated from the same object. This evidence comes in the form of correspondence cues: shared properties of two or more signals that suggest they share a common source. The most obvious correspondence cues are spatial (Slutsky & Recanzone, 2001; Lewald & Guski, 2003, Spence, 2013b) and temporal (Lewald & Guski, 2003; Magnotti, Ma, & Beauchamp, 2013; Parise, Spence, & Ernst, 2012; Slutsky & Recanzone, 2000; van Wassenhove, Grant, & Poeppel, 2007) coincidence, as predicted by the cell-recording work of Stein and colleagues (Meredith & Stein, 1983, 1985; Stein, Huneycutt, & Meredith, 1988). However, higher-order cues such as semantic (Adam & Noppeney, 2010; Laurienti, Kraft, Maldjian, Burdette, & Wallace, 2004), metaphoric (Parise & Spence, 2009; Sadaghiani, Maier, & Noppeney, 2009), or even learned arbitrary congruence (Altieri, Stevenson, Wallace, & Wenger, 2015; Ernst, 2007; Zweig, Suzuki, & Grabowecky, 2014) can also result in integration. One popular approach towards investigating multisensory integration is through illusions that exploit these correspondence cues. By providing strong evidence that they belong together, a researcher may force an observer to integrate signals that provide conflicting information. The specific conditions under which this effect occurs, and the nature of the resulting fused percept, can provide significant insight into the processes underlying multisensory integration.

The ventriloquist paradigm is perhaps the best-known example of this approach. The effect is named after ventriloquists' uncanny ability to produce the illusion of a voice emitting from the mouth of a dummy, and relies on the fact that spatial information is generally far more reliable in the visual modality than in auditory (we can see where things are far better than we can locate them by hearing). Presented simultaneously with a sound, a visual stimulus that is slightly displaced in space 'captures' the auditory signal, causing it to seem as though the sound originated closer to the visual stimulus than its true location. In a research setting, the effect was initially found to be most useful when the conflict between the signals was sufficiently small as to be imperceptible, as this may be used to determine the relative weighting the observer's sensory system gives to each signal (Ernst & Banks, 2002; Ernst, 2006; Jacobs, 1999). Alais and Burr (2004) presented participants with auditory and visual stimuli that were spatially separated by an undetectable amount and asked them to report their estimate of the combined audiovisual object's location. The reliability of the visual signal was changed between trials by adding varying levels of blur. Participants were found to combine the audiovisual signals in a statistically optimal way, weighting each modality by its relative reliability in line with the predictions of a maximum likelihood estimation (MLE) model (discussed in more detail in the following chapter). As the visual stimulus became more blurred, more weight was given to the auditory signal in the final estimate—in the most blurred cases, a 'reverse ventriloquist' effect was apparent, with the joint estimate lying closer to the true auditory location. Importantly, the combined estimates were also more reliable than unimodal localisation alone, demonstrating the benefits of multisensory integration for spatial localisation even when the cues are slightly in conflict.

This optimal weighting of integrated multisensory cues under a "forced-fusion" setting (i.e. the correspondence cues are sufficiently strong that signals are fully integrated) is not unique to audiovisual stimuli. Indeed, prior to the above study, statistically optimal integration was demonstrated for visual-haptic size judgement. Ernst & Banks (2002) presented participants with simulated visual and haptic objects, again with different levels of visual reliability, whose size differed between the modalities by an unnoticeable amount. Participants were asked to use the combined visual and touch information to judge the size of the virtual object. The contribution of the signals was again modulated by their relative reliabilities, and the combined estimate more reliable than unisensory judgements alone. Similar effects have been shown for visual-vestibular (Fetsch, Pouget, DeAngelis, & Angelaki, 2012), visual-proprioceptive (van Beers, Sittig, & van der Gon, 1999), and trimodal audio-visual-tactile (Bresciani, Dammeier, & Ernst, 2008) cue combinations. However, it is also becoming increasingly clear that humans do not perform optimally in all cases, instead often assigning more weight to the sensory modality that is most reliable for estimating a given property in everyday life (Meijer & Noppeney, 2018). This includes visual overweighting in spatial localisation tasks (Battaglia, Jacobs, & Aslin, 2003), auditory overweighting for temporal judgements (Burr, Banks, & Morrone, 2009; Maiworm & Röder, 2011), haptic overweighting during slant discrimination (Rosas, Wagemans, Ernst, & Wichmann, 2005), and vestibular overweighting for self-motion judgements (Butler, Smith, Campos, & Bülthoff, 2010; Fetsch, Turner, DeAngelis, & Angelaki, 2009). One explanation for these findings is that observers also consider prior knowledge about the reliability of a sensory modality when assigning it a weight, even if this leads to a mathematically sub-optimal response.

Prior information, as well as other complexities of real-world multisensory integration, are not well accounted for by the forced-fusion model. Most signals received by the brain at any one time originate from different sources, so integrating them all under the assumption that they are produced by one object would be disastrous. Koerding et al. (2007) recognised this problem and developed a Bayesian model that allows for the optimal reliability-weighted integration described above while also accounting for the presence of prior information and the possibility of stimuli that do not share a common source. This model, outlined in detail in the following chapter, acknowledges that the sensory uncertainty about the individual properties of a pair of signals is accompanied by further uncertainty about their underlying causal structure. Therefore, two competing possibilities are modelled simultaneously: either the signals were produced by the same object and should be integrated, weighted by their respective reliabilities, or they were produced by different objects and should be processed separately. The relative probability of these two models is then evaluated, and a response given based upon the outcome and the observer's final decision strategy (see Wozny, Beierholm & Shams, 2010, and Chapter 2 of this thesis, for an account of possible decision strategies). The BCI model accounts for responses to ventriloquist stimuli that are presented at a wide range of spatial disparities.

The sound-induced flash illusion is also commonly used to investigate the properties of multisensory perception. First described by Shams, Kamitani, and Shimojo (2000), this illusion provides an analogue of the ventriloquist effect in the temporal domain, where auditory perception dominates (we can detect *when* something happened far better by hearing than by vision). It can be broadly broken down into two types: fission and fusion (Andersen, Tiippana, & Sams, 2004). In the fission illusion, participants are presented with a brief visual stimulus accompanied by two beeps. If these beeps are spaced sufficiently close together in time, participants frequently report seeing two flashes. A greater number of beeps may lead to the perception of more flashes, but the effect becomes weaker as the number of beeps—or the time between them—increases. Fusion stimuli create the opposite effect. Multiple flashes, presented concurrently with a single beep, may be fused such that only a single flash is perceived. Interestingly, by configuring the stimuli such that the auditory stimulus was less reliable (i.e. presenting it at near-threshold hearing levels), Anderson et al. (2004) were also instead able to influence the perceived number of beeps using visual stimuli. This is analogous to the "reverse-ventriloquist" effect in the spatial domain and provides another neat demonstration of the role of stimulus reliability in optimal cue integration. Participants' responses to the sound-induced flash illusion have been successfully modelled using the same BCI approach as that applied to the ventriloquist paradigm (Shams, Ma, & Beierholm, 2005), providing evidence that a Bayesian framework may be useful in describing multisensory integration as a general process.

A final illusion that is regularly applied to the investigation of multisensory integration is concerned with speech perception. In the McGurk illusion (McGurk & MacDonald, 1976; Tiippana, 2014), participants are presented with a video of a person articulating a syllable, accompanied by the sound of a different syllable. Depending on the specific combination of syllables used, an entirely new percept can be created; in the original form of the illusion, a visual [ga] was coupled with an auditory [ba], creating the perception of [da]. Intriguingly, when this combination was reversed (visual [ba] with auditory [ga]), the resulting percept was instead often a sequential combination such as [baga]. Finally, an auditory [ba] with visual [da] was perceived as [da]. As with other multisensory illusions, the McGurk effect may be modulated by stimulus reliability (Campbell & Massaro, 1997; MacDonald, Andersen, Bachmann, 2000; Thomas & Jordan, 2002). However, the wide array of percepts generated by the many combinations of incongruent signals possible in this illusion have proven difficult to model. Magnotti and Beauchamp (2017) have recently proposed a causal inference model of audiovisual speech perception with some success, but the model constrains observers' responses to a selection of single-syllable possibilities, limiting its ability to account for the more nuanced percepts described by McGurk and MacDonald (1976). That efforts to fully account for this illusion within the BCI framework are still ongoing is testament to the relative complexity of multisensory speech integration.

These illusions, as well as an array of others that include the stream/bounce illusion (Sekuler, Sekuler, & Lau, 1997), size-weight illusion (Murray, Ellis, Bandomir, & Ross, 1999), and rubber-hand illusion (Botvinick & Cohen, 1998), are invaluable tools for investigating the mechanisms underlying multisensory integration. We are not constrained to these approaches, however. Multisensory enhancement is also an extremely useful metric for determining the influence of one stimulus modality on another, and is perhaps the behavioural outcome that most clearly relates to the super-additive neuronal activity discussed earlier (Sperdin, Cappe, Foxe, & Murray, 2009; Stein, Huneycutt, & Meredith, 1988). Multisensory enhancement can manifest in several ways. One of the most robust, and easiest to measure, is reaction time (RT). Participants respond to congruent multisensory stimuli significantly faster than to unisensory stimuli, regardless of whether the task involves simple detection (Colonius, 1988; Forster, Cavina-Pratesi, Aglioti, & Berlucchi, 2002; Miller, 1982) or some form of decision about the stimulus (Giard & Peronnet, 1999; Hairston, Laurienti, Mishra, Burdette, & Wallace, 2003; Harrington & Peck, 1998; Hershenson, 1962; Laurienti, Burdette, Maldjian, & Wallace, 2006). Diederich and Colonius (2004) even demonstrated increased RT enhancement in response to trimodal (auditory/visual/somatosensory) stimuli. Incongruent multisensory stimuli, conversely, can negatively impact reaction times (Cowan & Barron, 1987; Laurienti, Kraft, Maldjian, Burdette, & Wallace, 2004; Sinnett, Soto-Faraco, & Spence, 2008).

In addition to improving reaction times, multisensory stimuli can also lower detection thresholds. Lippert, Logothetis, and Kayser (2007), for example, showed that a simultaneous auditory stimulus can improve visual detection (see also Adam & Noppeney, 2014; Frassinetti, Bolognini, & Làdavas, 2002; Noesselt, Bergmann, Hake, Heinze, & Fendrich, 2008). Lovelace, Stein, and Wallace (2003) found a similar effect in the opposite direction, with visual stimuli improving auditory detection, and tactile enhancement of auditory detection thresholds has also been shown (Gillmesiter & Eimer, 2007).

Finally, multisensory integration can enhance comprehension of speech stimuli, especially in situations where the auditory stimulus is degraded. Sumby and Pollack (1954) demonstrated that, under noisy conditions, the addition of congruent visual information significantly improved participants' comprehension of spoken words. This has proven to be robust and highly-replicable (Erber, 1975; Grant & Seitz, 2000; Lachs & Pisoni, 2004; Macleod & Summerfield, 1986; McGettigan et al., 2012; Schwartz, Berthommier, & Savariaux, 2004). The mechanisms underlying this effect are not entirely clear, but numerous attempts have been made to discern the specific features of the added visual information that lead to enhanced comprehension. Schwartz et al. (2004), for example, were able to demonstrate that the enhancement cannot be attributed simply to an improvement in detectability, as a neutral stimulus (red square) indicating the onset of a target syllable did not improve comprehension in the context of background noise. Grant and Seitz (2000; also Grant, Walden, & Seitz, 1998) argue that viewing a speaker provides complementary information about specific features of speech sounds, such as the point of articulation, that are easily disrupted by the presence of auditory noise. Interestingly, however, Reisburg, McLean, and Goldfield (1987) found that conceptually difficult information is more readily understood when the observer can both see and hear the talker, even under silent conditions. This suggests that there is also some benefit to presenting entirely redundant speech information via multiple sensory modalities, perhaps pointing towards the involvement of higher-level attentional mechanisms (Campbell, 2008).

Cognitive factors modulate multisensory integration

There is some evidence that multisensory integration may be subject to top-down influence. Aside from the existence of the higher-level correspondence cues discussed earlier, research has also revealed interactions with attention, context, and task demands. In the case of multisensory speech processing, higher-order processes clearly play an important role. Alsius, Navarra, Campbell, and Soto-Faraco (2005) found that the strength of the McGurk illusion was much diminished under a high attentional load induced by embedding the illusory stimuli randomly within a dual-task paradigm. Alsius, Navarra, & Soto-Faraco (2007) demonstrated a similarly reduced effect when participants directed their attention towards a difficult tactile judgement task during the presentation of the McGurk stimuli. Nahorna, Berthommier, and Schwartz (2012; 2015) and Gau and Noppeney (2015) were also able to modulate the McGurk effect by manipulating the presentation context, suggesting an effect of prior expectation of congruence on the multisensory integration of speech stimuli. Finally, the strength of the McGurk illusion has been shown to differ considerably depending on the response required (Colin, Radeau, & Deltenre, 2005).

The picture is less clear for non-speech stimuli. Bertelson, Vroomen, de Gelder, & Driver (2000) and Vroomen, Bertelson, and de Gelder (2001) presented convincing evidence that the ventriloquist effect is modulated by neither endogenous nor exogenous attention (respectively). Busse, Roberts, Crist, Weissmann, and Woldorff (2005) found similar behavioural results, but did report differences in cortical activation (using both EEG and fMRI) between attentional conditions. Two studies that measured multisensory cue weighting report directly conflicting results: Vercillo and Gori (2015) found that selectively attending to the auditory stimulus during an auditory-haptic localisation task significantly increased both the precision and relative weighting of the sound signal, while Helbig and Ernst (2008) were unable to manipulate visual-haptic cue weighting using a similar approach. Multisensory integration of motion, conversely, appears to be strongly impacted by task demands and modality-specific attention (Donohue, Green, & Woldorff, 2015; Oruc et al., 2008). Finally, Lippert, Logothetis, and Kayser (2007) neatly demonstrated an effect of learned association—

a heavily top-down process—on the auditory enhancement of visual detection thresholds. A simultaneous sound that was directly informative about the onset time of a low-contrast visual stimulus only improved participants' detection rates when presented in a context that gave clues to its usefulness. The degree to which multisensory integration is modulated by top-down processes is thus not fully understood and seems to be dependent on features of both stimuli and task, a point that will remain relevant as I explore the effects of healthy ageing on these integration processes.

Healthy ageing of the brain

Having summarised the current state of research into multisensory integration, in terms of perception and behaviour as well as the neural systems supporting these, I will now give a brief overview of the known effects of ageing on the human brain and sensory systems. It is important to clarify that I am explicitly interested in healthy ageing; that is, changes that occur to the structure and function of these systems over time in the absence of diseases such as dementia or stroke.

Physiological changes

Structural decline

One of the most consistent and (since the advent of neuroimaging) easily-measurable effects of chronological age on the brain is a loss of grey matter volume. Raz et al. (2005), for example, contacted older adults who had taken part in cross-sectional imaging research five years earlier and invited them for a follow-up scan. Manual demarcation of grey matter in each scan revealed significant volume reductions over time in a variety of regions including lateral prefrontal cortex, orbitofrontal cortex, inferior temporal cortex, and fusiform cortex. In a larger cross-sectional study, Good et al. (2001) used automated voxel-based morphometry to measure grey matter volume in 465 participants aged between 18 and 79 years old, and also found significant age-related reduction in grey matter volume across a number of regions. Although there is some variation in terms of estimated rates of tissue loss (a summary in Crivello, Tzourio-Mazoyer, Tzourio, & Mazoyer, 2014 reports rates between –1.4 and –4cm^3 per year), relative effects on specific regions (e.g. Crivello et al. report substantial atrophy in occipital cortex, while Raz et al., 2005, 2010, find this region to be largely unaffected), and gender differences (Good et al. found faster tissue loss in males, while Crivello et al. report the opposite; Raz et al., 2005, saw no effect of gender) a considerable number of large longitudinal and cross-sectional studies support the general conclusion that chronological ageing is closely associated with a loss of grey matter (Blatter et al., 1995; Crivello et al., 2014; Fjell et al., 2009; Giorgio et al., 2010; Good et al., 2001; Jernigan et al., 1991; Jernigan et al., 2001; Pfefferbaum et al., 1994; Raz, Ghisletta, Rodrigue, Kennedy, & Lindenberger, 2010; Raz et al., 2005; Raz et al., 2007; Salat, Kaye, & Jonowsky, 1999; Sowell et al., 2003; Sullivan, Rosenbloom, Serventi, & Pfefferbaum, 2004; Walhovd et al., 2005).

The mechanisms underlying grey matter atrophy are less clear (Fjell & Walhovd, 2010). For a considerable time it was believed that ageing results directly in the loss of neurons (e.g. Brody, 1955). However, methodological advances have since led to the emergence of a more nuanced picture. Terry, DeTeresa, & Hansen (1987), for example, used a combination of manual and automatic approaches to assess neuronal quantity, size, and density in a sample of post-mortem brains from adults aged 24 to 100 years. The authors found that the raw number of cells was in fact similar across age groups, but that the average size of the neurons was reduced in older adults. A review by Peters, Morrison, Rosene, Hyman, and Peters (1998) agreed with this assessment, concluding that little evidence exists to suggest that neuronal count declines with age. Instead, it seems that loss of grey matter volume may be attributed, at least in part, to the decline in neuronal size described by Terry and colleagues (Esiri, 2007; Jacobs, Driscoll, & Schall, 1997).

The volume of white matter may also decline with age, but the differences are less pronounced or consistent between studies. Some report little or no age-related change in white matter volume (Blatter et al., 1995; Good et al., 2001; Jernigan et al., 1991; Pfefferbaum et al., 1994; Sullivan et al., 2004), while others report either localised or general white-matter loss in older adults (Giorgio et al., 2010; Guttman et al., 1998; Jernigan et al., 2001; Pagani, Agosta, Rocca, Caputo, & Filippi, 2007; Raz et al., 1997; Raz et al., 2005; Salat et al., 1999; Walhovd et al., 2005). These discrepancies are likely due in part to methodological differences in the approach towards quantifying white matter volume, with each using a unique combination of manual and automatic classification of images of differing types (e.g. T1-weighted versus diffusion tensor images). However, such methodological diversity also exists in the measurement of cortical grey matter, and yet the conclusion in that case is unanimous; white matter volume is thus either harder to quantify, or less robustly affected by age.

There is little doubt that the integrity of white matter tissue, if not its volume, is negatively affected by advancing age (Gunning-Dixon, Brickman, Cheng, and Alexopoulos, 2009; Liu et al., 2017). Myelin sheaths have been shown to split or become deformed with age (Peters, 2002), an effect that is compounded by the concurrent degeneration of myelin-replenishing oligodendrocytes (Kohama, Rosene, & Sherman, 2012). As myelination is critical for efficient conduction along the neuronal axon, the functional outcome of this degradation is reduced conduction speed and efficiency, and therefore impaired processing speeds (Liu et al., 2017; Peters, 2002; Kochunov et al., 2010).

Cortical vasculature

Underlying many of these neurodegenerative processes is a decline in cerebrovascular health (Kalaria, 2010). The brain, as living tissue, is dependent upon a supply of oxygen and other nutrients that is both reliable and reactive—able to adjust to meet demand. Increasing age is known to be a significant risk factor for cardiovascular problems that interfere with this supply (D'Esposito, Deouell, & Gazzaley, 2003; Grady and Garrett, 2014). Cardiovascular health has therefore become a prominent target for interventions that aim to address age-related neurodegeneration, and several studies have demonstrated successful reversal of brain ageing markers, including tissue loss, through the introduction of exercise regimes (Colcombe et al., 2006; Gordon et al., 2008; Hillman, Erickson, & Kramer, 2008; Marks et al., 2007). A practical implication of age-related changes in cerebrovascular health for neuroimaging research, discussed in more detail in Chapter 2, is that they can make age differences in neural activity more difficult to quantify: functional MRI relies on measures of blood flow and oxygenation as indices of neural activation, so results may be confounded if these are altered due to other physiological factors (D'Esposito et al., 2003; Tsvetanov et al., 2015).

Functional activity

Finally, brain ageing has been associated with considerable changes in patterns of task-related neural activity across a wide variety of stimuli and cognitive tasks. These differences often manifest as an overall increase in activity and/or shift in activity towards other regions. Cabeza (2002), for example, outlines a hemispheric-asymmetry reduction in older adults (HAROLD) model that describes a tendency for prefrontal activity in older adults to be less lateralised than in younger age groups. Davis, Dennis, Daselaar, Fleck, & Cabeza's (2008) posterior-anterior shift in ageing (PASA) theory describes a similar tendency for prefrontal activity to increase (sometimes accompanied by an occipital decrease) in older adults. Increased activity is perhaps contrary to what might be expected in a population with decreased cortical mass, leading researchers to theorise that it may represent compensatory activity. For example, Park and Reuter-Lorenz (2009) proposed the scaffolding theory of ageing and cognition (STAC), arguing that older adults continually and dynamically recruit extra regions to compensate for deterioration in others. This suggestion is similar to older ideas of cognitive flexibility or reserve (Barulli & Stern. 2013; Stern, 2002), which suggest that older adults who have greater cognitive adaptability (as a result of environment and life experience) are better protected against the deleterious effects of ageing and neuropathology because they can potentially call on multiple systems to solve the same problem. Supporting STAC and related compensation theories is research demonstrating that older adults who show these activity changes perform better on cognitive tasks than those who do not (e.g. Cabeza, Anderson, Locantore, & McIntosh, 2002; Cherry et al., 2005; Fera et al., 2005; Reuter-Lorenz et al., 2001). Furthermore, in a longitudinal study of 40 older adults, Persson et al. (2006) discovered that those who showed the greatest loss of hippocampal grey matter also displayed an increase in prefrontal cortex activation. Park and Reuter-Lorenz (2009) interpret this as evidence of higher compensatory activity in the older adults who most needed it.

As acknowledged by Persson and colleagues (2009), however, there is another possible interpretation of these results. Age-related increases or regional changes in neural activity in response to a task may instead simply reflect noisy or inefficient activations. As Morcom and Johnson (2015) explain, it is extremely difficult to differentiate between inefficient and compensatory activity, particularly using standard univariate analysis methods, and the evidence that currently exists is inconclusive. One approach is to employ multivariate analyses that allow for interrogation of the type of information that is represented within a region: do these areas with increased activation also hold more information about the task or stimulus? A recent study that applied this approach to two fMRI datasets (Morcom, Henson, & Cam-CAN, 2018) concluded that older adults' greater prefrontal activations to memory tasks did not coincide with an increase in the amount of relevant information held in this area, suggesting that they were not in this case compensatory. This approach towards differentiating between inefficient and compensatory activation is still in its infancy, however, and further work in a greater variety of tasks is required before wider conclusions can be drawn.

Changes in behavioural, cognitive, and sensory function

The changes to the physical structure and function of the brain that occur with healthy ageing are accompanied by decline in a variety of sensory, cognitive, and behavioural abilities. These effects are strikingly inconsistent across task domains, however, which may be due in some part to differential rates of physiological decline in the supporting regions. Steffener, Brickman, Habeck, Salthouse, & Stern (2013) found that cerebral blood flow and grey matter volume predicted different aspects of cognitive performance, and that the two combined accounted for almost all age-related differences in their tests of memory, speed, attention, and abstract reasoning. Similarly, a series of studies by Raz and colleagues (Gunning-Dixon & Raz, 2003; Head, Kennedy, Rodrigue, & Raz, 2009; Kennedy, Rodrigue, Head, Gunning-Dixon, & Raz, 2009) revealed differential links between grey and white matter decline (in various regions) and multiple tests of cognitive performance. Here I will briefly outline some of the functions that are most commonly shown to be impaired by healthy ageing.

Reaction times

Perhaps the most consistently replicable behavioural result of healthy ageing is increased reaction times, an effect that is apparent to some degree across most tasks (Der &

Deary, 2006; Verhaegen & Cerella, 2008). Some of this may be attributed to declines in motor speed (Ebaid, Crewther, MacCalman, Brown, and Crewther, 2017; Wilson, Bienias, Evans, & Bennett, 2004), but substantial differences in the size of the effect between tasks suggest that it also indexes other important processes underlying cognitive performance (Salthouse, 1996; Sliwinkski & Hall, 1998). Indeed, reaction times (RTs) are commonly used to compare performance between younger and older adults. However, it is important to note that older adults also tend to set their response criteria differently, prioritising accuracy to a greater degree than younger adults even in explicit tests of reaction times (Rabbitt, 1979; Salthouse, 1979; Starns & Ratcliffe, 2010). Any study using age differences in reaction times to index underlying cognitive processes should therefore aim to account for confounding effects of motor speed and speed/accuracy tradeoff. When summary RT measures are used, interactions between age and task conditions should also be treated with caution as they can be difficult to interpret, particularly when a large main effect of age is present (Wagenmakers, Krypotos, Criss, & Iverson, 2011).

Memory

Impaired memory is probably the most stereotypically recognisable impact of cognitive ageing. The degree of the effect is highly dependent on the type of memory and task, however. Procedural memories that were formed early in life and are regularly retrieved are generally well preserved, as is semantic memory, while explicit recall of recent events can be far weaker in older adults (Zacks, Hasher, & Li, 2000). Older participants generally perform far worse in tasks that require free or cued recall compared to those that involve selecting a previously encountered stimulus from a new set (Grady & Craik, 2000). The literature exploring the effects of ageing on memory is vast, and further discussion is beyond both my expertise and the scope of this thesis; I instead refer the interested reader to the excellent reviews that have been written on the topic (Cohen, 2014; Grady & Craik, 2000; Luo & Craik, 2008; Zacks, Hasher, & Li, 2000).

Selective attention and executive functions

Executive functions, including abilities such as planning, problem solving, attentional control, response inhibition and evaluation, and similar high-level cognitive functions, are frequently found to be affected by ageing. I also include selective attention here; while it is not generally considered together with executive functions, it relies on/overlaps with attentional control and response inhibition so heavily that these are often measured by the same tasks (see e.g. Desimone & Duncan, 1995). The interplay between these functions and healthy ageing is another wide and complex area spanning decades of research. However, as this area is important as a reference point to explain findings later in this thesis (activation of cognitive control/inhibition regions in Chapters 4 and 5), I have attempted to give an overview of the most relevant findings.

Selective attention and executive functions are crucial to performance in many of the tasks used to assess multisensory integration. Indeed, many multisensory integration problems may easily be recast in terms of these processes. A ventriloquist or sound-induced flash illusion task, for example, usually requires selective attention to one modality and inhibition of responses to the other. Added visual information in a speech-in-noise task may be considered as aiding selective attention to the speech signal.

The Stroop task provides an excellent case study of age differences in these functions. In the basic form of this paradigm, participants view the names of various colours displayed in fonts of non-matching colours. The task is to state the colour each word is displayed in, as quickly as possible, while ignoring the word itself. This is difficult because the written colour interferes with the required response; it is a classic executive function task that relies heavily on selective attention and response inhibition. The size of the effect is generally quantified as the difference in response times to Stroop stimuli compared with an interference-free control condition.

It has been well established by multiple large-scale studies that older adults show a greater Stroop effect than their younger counterparts (e.g. Bugg, DeLosh, Davalos, & Davis, 2007; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2006). This correlates with age-related neural activity changes in a network of attentional control regions including anterior cingulate, dorsolateral prefrontal cortex, inferior frontal gyrus, and anterior insula, suggesting that this system is changed or impaired in older adults (Langenecker, Nielson, and Rao, 2004; Milham et al., 2002; Mohtasib et al., 2012; Schulte et al., 2011; Zysset et al., 2007). Age differences in this task are also apparent in an auditory equivalent (Sommers & Danielson, 1999), and present even after accounting for various types of reaction time (Bugg et al., 2007), supporting the argument that the effect is truly indexing some impairment in attentional and executive function processes.

This behavioural effect is not confined to the Stroop task: older adults are more affected by task-irrelevant distractors across a variety of stimulus types and sensory modalities. Gazzaley, Cooney, Rissman, & D'Esposito (2005) found that older adults were less able to recall recently viewed stimuli when they were interspersed with irrelevant distractors, and that the strength of this deficit correlated with impairments in working memory. Further, in a global/local visual selection task that avoided any memory confounds, Tsvetanov, Mevorach, Allen, & Humphreys (2013) showed that older adults were less capable of ignoring highly salient visual distractors. Similar effects have been found in many selective attention/executive function tasks (e.g. Helfer & Freyman, 2008; Mahoney, Verghese, Goldin, Lipton, & Holtzer, 2010; Presacco, Simon, & Anderson, 2016; Schmitz, Cheng, & De Rosa, 2010; see below the section on ageing and multisensory integration for examples of cross-sensory distraction).

Age differences in the neural network underlying executive functions are also not exclusive to studies of the Stroop task. A meta-analysis by Turner & Spreng (2012) revealed age differences in activation of regions including inferior frontal gyrus, dorsolateral prefrontal cortex, and insula in response to tasks as diverse as flanker, visual search, and go/no-go paradigms. Interestingly, however, Zhu, Zacks, & Slade (2010) showed that older adults' BOLD responses during a rapid arrow flanker task most closely matched those of younger adults on trials where both groups answered successfully, suggesting that at least some of the age-related activation differences observed in other studies may be the result of increased error processing in the older age group.

Sensory acuity

Central sensory processing is also affected to some degree by healthy ageing, but to a far lesser extent. A primary driver of this is loss of peripheral sensory acuity (Gates & Mills, 2005; Klein, Klein, Linton, & De Mets, 1991). However, age-related changes in sensory processing have been observed even in the absence of peripheral impairments. Some examples in the auditory domain are discussed in the following section on multisensory integration and ageing: older adults are less able to comprehend speech presented in a noisy environment (Dubno, Dirks, & Morgan, 1984; Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004; Presacco, Simon, & Anderson, 2016; Wong et al., 2009) and have less precise representations of auditory temporal structure in general (Fitzgibbons & Gordon-Salent, 2004; Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013). Similar temporal impairments have also been observed for visual stimuli (Eriksen, 1970). Visual perception of motion (Habak & Faubert, 2000; Allard, Lagacé, & Faubert, 2013), symmetry (Herbert, Overbury, Singh, & Faubert (2002), and stereoscopic depth perception (Wright & Wormald, 1992) have also been shown to decline with age.

Preserved functions

Finally, it is worth noting that numerous cognitive functions show little or no decline with age. In particular, scores on tests that in some way measure accumulated knowledge or experience improve or remain stable until the very latest stages of adult life (Salthouse 2010, 2012). These include, but are not limited to, vocabulary and verbal fluency (Tomer & Levin, 1993; Verhaegen, 2003), procedural memory (Churchill, Stanis, Press, Kushelev, & Greenough, 2003), and "wisdom-related" knowledge (Smith & Baltes, 1990; Staudinger, Smith, & Baltes, 1992).

Conclusion

To conclude, the normal, healthy process of ageing has a wide range of physiological outcomes, which may interact with measures of brain activity, as well as behavioural, cognitive, and sensory outcomes, which may change or interact with multisensory integration. In order to best understand them, it is therefore important to consider any observed age differences in multisensory integration in the context of these various other effects.

Ageing and multisensory integration

In the final section of this chapter I review our current understanding of the effects of healthy ageing on multisensory integration. I also include some studies of crossmodal attention as these often address similar questions (as noted above, many tasks that measure multisensory integration are effectively selective attention paradigms). The approaches used are as diverse as their findings, but I have grouped them into general themes.

Reaction time enhancement

In terms of multisensory enhancement of reaction times, arguably the simplest measure of multisensory integration, healthy older adults consistently perform at least as well as younger controls. Laurienti, Burdette, Maldjian, & Wallace (2006) used a speeded decision task to assess multisensory enhancement in younger and older adults. Participants were instructed to indicate the colour of a stimulus based on auditory, visual, or congruent audiovisual cues. Deconstruction of the data using a race model analysis revealed that both age groups benefited from the presentation of bisensory cues. Indeed, in this case the increase in reaction time performance was greater in the older adults. This age-related improvement is not always replicated, but similar studies using a variety of simple detection (Diaconescu, Hasher, & McIntosh, 2013; Mahoney, Li, Oh-Park, Verghese, & Holtzer, 2011; Peiffer, Mozolic, Hugenschmidt, and Laurienti, 2007;) and discrimination (Diaconescu et al., 2013; Diederich, Colonius, & Schomburg, 2008; Hugenschmidt, Mozolic, & Laurienti, 2009; Wu, Yang, Gao, & Kimura, 2012; Zou, Chau, Ting, and Chan, 2017) tasks have found at least equal audiovisual enhancement of reaction times in younger and older adults (though see Stephen, Knoefel, Adair, Hart, & Aine, 2010, where an effect appears to be present but was non-significant). A series of studies by Mahoney, Holtzer and colleagues demonstrated that this preservation of multisensory enhancement with age is not restricted to audiovisual stimuli and also holds true for visual-somatosensory (Mahoney, Dumas, & Holtzer, 2015; Mahoney, Holtzer, & Verghese, 2014; Mahoney et al., 2011; Mahoney, Wang, Dumas, & Holtzer, 2014) and auditory-somatosensory (Mahoney et al., 2011) stimulus combinations.

Impaired temporal causal inference

It is when stimuli become in some way incongruent, and observers must therefore make judgements about their underlying causal structure before giving a response, that age differences begin to emerge. For example, a considerable number of studies have demonstrated that older adults are susceptible to the sound-induced flash illusion across a far wider range of stimulus onset asynchronies than younger adults (Chan, Connolly, & Setti, 2018; DeLoss & Andersen, 2015a, 2015b; DeLoss, Pierce, & Andersen, 2013; McGovern, Roudaia, Stapleton, McGinnity, & Newell, 2014; Parker and Robinson, 2018, Setti et al., 2014; Setti, Burke, Kenny, & Newell, 2011; Stapleton, Setti, Doheny, Kenny, & Newell, 2013), suggesting that they are less able to segregate temporally distinct events. This may be at least partly due to impaired multisensory temporal perception (the ability to determine the relative timings of events perceived by different senses): older adults have been shown to be less able to distinguish the temporal order of audiovisual signals presented in rapid succession (Bedard and Barnett-Cowan, 2015; de Boer-Schellekens and Vroomen, 2013; Setti, Finnigan, et al., 2011; but see Fiacconi, Harvey, Sekuler, and Bennett, 2013), and Setti, Stapleton, et al. (2014) demonstrated that training on such temporal-order tasks could reduce older adults' susceptibility to the sound-induced flash illusion. Setti, Finnigan, et al. (2011) investigated the neural mechanisms underlying age differences in this temporal-order judgement task via an ERP study, finding reduced response amplitudes to the second stimulus in older adults selectively for conditions in which they were less accurate. The authors conclude that it is the processing of this second stimulus in the sequence that is impaired in older adults.

Age also leads to an increased susceptibility to the stream/bounce effect, another timing-critical audiovisual illusion in which two visual stimuli are seen to converge and the probability of them being perceived as bouncing off each other may be increased by the presentation of a sound at the point of overlap. As with the sound-induced flash illusion, Bedard and Barnett-Cowan (2015) demonstrated that older adults are susceptible to this effect across a wider range of audiovisual asynchronies, again suggesting that they are less able to segregate temporally distinct stimuli, though it must be noted that this age difference is not especially robust and can be reversed by certain stimulus manipulations (Roudaia, Sekuler, Bennett, & Sekuler, 2013).

More explicit simultaneity-judgement tests have produced equivocal results. Bedard and Barnett-Cowan (2015) report that, in the same large sample that showed age-related differences in both temporal-order judgements and the stream/bounce illusion, age had no effect on explicit audiovisual synchrony judgements. Stevenson, Baum, Krueger, Newhouse, and Wallace (2017) and Chan, Pianta, & McKendrick (2014a) did find older adults to be more likely to report temporally distinct stimuli as occurring simultaneously, but these age differences were small.

The weight of evidence in these mixed findings does indicate that healthy ageing has some effect on the integration of multisensory stimuli across time, but the effect seems to be nuanced and can be difficult to account for. The direction of the findings suggests that older adults have a larger multisensory temporal binding window. That is, they are more likely to integrate stimuli that are more asynchronous than younger people. Characterised in terms of the BCI framework (Shams, Ma, & Beierholm, 2005), this could be explained by either a greater common-source prior (prior tendency to bind signals) or decreased sensory variance (reliability of timing estimates). The existence of research that also demonstrates age-related unisensory timing impairments (e.g. Fitzgibbons & Gordon-Salent, 2004; Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013) suggests the latter is more likely, but to my knowledge no results have yet been published that directly address this distinction. It must also be noted that tasks that have used explicit measures of integration (directly asking participants whether two signals matched) diverge in terms of findings from those using implicit measures (measuring an outcome of integration, as with illusions), suggesting that the mechanisms underlying these differ and are differentially affected by ageing.

Preserved spatial integration?

Research investigating similar interactions between age and multisensory integration in the spatial domain is far scarcer. To my knowledge, nobody has yet tested the effects of ageing on susceptibility to the ventriloquist effect. Dobreva, O'Neill, and Paige (2012) found that older adults were slightly less accurate at reporting the location of unimodal and bimodal targets, and that the visual component of bimodal stimuli had a stronger influence for older adults, but this was based only on comparisons between unisensory and spatially congruent bisensory trials. Other approaches have produced mixed results, with spatial manipulations sometimes failing to affect responses in either age group. DeLoss and Andersen (2015a) attempted to manipulate the strength of the sound-induced flash illusion by increasing the spatial separation between the auditory and visual signals, but neither younger nor older adults were influenced by this disparity. Similarly, Mahoney, Wang, et al. (2014) showed that introducing spatial disparity between stimuli had no influence on multisensory reaction time enhancement in either age group. Campbell, Al-Aidroos, Fatt, Pratt, and Hasher (2010) did create trajectory deviations by introducing a visual distractor when participants made saccades towards a multisensory target, but also found no effect of age, while Barrett and Newell (2015) showed that a spatially incongruent distractor impaired reaction times similarly in both age groups.

Others have shown age-related effects of stimulus location and disparity, however. Poliakoff, Ashworth, Lowe, & Spence (2005) found, in a visual-tactile paradigm, that the presence of a cross-modal distractor led to decreased performance in young (19-25), young-old (65-72) and old-old (76-92) adults, but only in the younger two age groups was this modulated by the spatial disparity between signals. In contrast, Diederich, Colonius, and Schomburg (2008) found that older adults benefited less from audiovisual signals presented at spatially disparate locations.

Weighting of multisensory cues

An important question to answer on the path to understanding how age affects multisensory integration is whether older adults assign the same relative weights to integrated cues as younger adults, who generally perform near optimally (see earlier section on forced-fusion integration). To my knowledge only three studies have thus far attempted this. Brooks, Anderson, Roach, McGraw, and McKendrick (2015) assessed relative weighting of auditory and visual cues in a temporal rate perception paradigm and found that older and younger adults did assign similar weights to the cues, but this only resulted in a (small) performance increase in the younger group. A visual-haptic verticality judgement task reported by Braem, Honoré, Rousseaux, Saj, and Coello (2014) also showed optimal cue weighting in younger and healthy older adults. However, in a navigation task that required integration of visual and self-motion cues, Bates and Wolbers (2014) found that older adults underweighted the visual information relative to an MLE-optimal prediction.

Multisensory speech integration

An area of very active investigation is the interaction between ageing and audiovisual speech integration, especially in the context of competing background noise, which has been shown to have a greater impact on comprehension in older age groups (Dubno, Dirks, & Morgan, 1984; Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004; Wong et al., 2009). As the presence of lip movement information is known to substantially improve speech detection (Bernstein, Auer, & Takayanagi, 2004) and comprehension (MacLeod & Summerfield, 1986; Schwartz et al., 2004; Sumby & Pollack, 1954) in noisy environments, multisensory integration has the potential to “offer a fountain of youth for older ears” (Winneke and Phillips, 2011).

Tye-Murray, Spehar, Myerson, Hale, and Sommers (2016), and Tye-Murray, Sommers, Spehar, Myerson, and Hale (2010), using build-a-sentence tests that required participants to identify target words in a sentence presented with background noise, found that intact visual information is indeed useful for older adults, and that it improves sentence comprehension to a similar degree in both age groups. Other work, including an earlier consonant discrimination study by the same group (Sommers, Tye-Murray, & Spehar, 2005), and sentence repetition (Helfer, 1998) and word recognition (Anderson Gosselin & Gagné, 2011) tests by others, confirm this similarity. Age differences do emerge when the task becomes more challenging, however. Dey and Sommers (2015) report reduced visual enhancement of speech comprehension in older adults, but only when the task was to identify lexically hard words (those with many close alternatives). Similarly, in the context of extremely challenging listening environments (SNR of -12 and lower), Stevenson et al. (2015) found that younger participants exhibited far more multisensory benefit in whole-word recognition. Performance was comparable across age groups for a less-challenging but otherwise equivalent phoneme recognition task.

Regardless of context or task, there is considerable evidence that older adults rely more on the visual component of a multisensory speech signal than do younger adults. Tye-Murray, Sommers, et al. (2010), for example, found that an extremely-low-contrast video stimulus facilitated comprehension less in the older age group. Gordon and Allen (2007) report even more extreme age differences, with older adults receiving almost no benefit from blurred visual signals that improved comprehension substantially in the younger age group (though see also Legault, Gagné, Rhoualem, and Anderson Gosselin, 2010). Maguinness, Setti, Burke, Kenny, and Newell (2011) found that pixilating the visual stimulus can impair older adults' audiovisual speech comprehension even when the auditory stimulus is intact and clear, indicating that they are also reliant on visual information to support their understanding in non-challenging environments. Furthermore, Winneke and Phillips (2011) showed that, despite equivalent levels of behavioural multisensory enhancement in an audiovisual speech recognition task, older adults showed earlier and larger multisensory ERP responses consistent with a greater influence of visual cues.

This idea of visual dominance in older adults' multisensory speech perception is supported by multiple studies that utilised McGurk-type paradigms. By providing participants with a range of response options it is possible to use the McGurk stimuli to measure sensory biases, as some responses are more compatible with the auditory signal and others with the visual. Cienkowski and Carney (2002) showed that although younger and older adults are similarly susceptible to the McGurk illusion, older adults were overall more likely to give visually-biased responses (responses that were more compatible with the visual stimulus than the auditory). More extensive work has since showed either increased susceptibility to the McGurk illusion older adults or greater visual bias in responses (Huyse, Leybaert, and Berthommier, 2014; Sekiyama, Soshi, & Sakamoto, 2014; Setti, Burke, et al., 2013; Stevenson et al., 2017).

Multisensory attentional interactions

A number of studies have attempted to assess whether attentional processes may interact with multisensory integration differently in younger and older adults. In some cases, older adults appear to be more susceptible to cross-sensory distraction. In a novel simulated driving study, Pitts and Sarter (2018) asked participants to detect unisensory or simultaneous bi-, and tri-sensory stimuli whilst maintaining a continuous driving position. Older participants were both less able to accurately detect the multisensory stimuli, and more likely to deviate in driving position following stimulation. Similarly, Mevorach, Spaniol, Soden, & Galea, (2016) showed that older adults were more impaired by the presence of visual distractors during a motor task. Guerreiro and Van Gerven (2011; also Guerreiro, Murphy, and Van Gerven, 2013) also found that older participants were more susceptible to visual distraction in an auditory *n*-back task, and vice versa, but later EEG and fMRI replications using similar stimuli (Guerreiro, Anguera, Mishra, Van Gerven, and Gazzaley, 2014; Guerreiro, Eck, Moerel, Evers, and Van Gerven; 2015) found no age differences in either behavioural or neural responses. Townsend, Adamo, & Haist (2006) also found no age-related behavioural differences in an fMRI focused-attention study, but do report that older adults displayed significantly increased frontoparietal activations in response to irrelevant stimuli. The authors suggest that this may be compensatory recruitment to adjust for reduced sensory inhibition.

Findings are also mixed for cross-modal cueing paradigms. Mahoney, Verghese, Dumas, Wang, and Holtzer (2012) report that older adults receive more reaction time benefit from multisensory cues that have a visual component than do younger adults. Meanwhile, Andrés, Parmentier, and Escera (2006) found that simple beep cues benefitted both age groups in a speeded visual decision task, but a more complex natural sound actually impaired older adults' reaction times. Others have found no age differences in multisensory cueing tasks. DeLoss et al. (2013) were able to modify susceptibility to the sound-induced flash illusion via cued modality-specific attention, but this affected both age groups comparably. Similarly, Guerreiro, Adam, Gerven, & Pascal, 2012 found age did not affect the impact of auditory or visual cues on reaction times to unisensory cues, and Campbell, Al-Aidroos, Fatt, Pratt, and Hasher (2010) showed both age groups were similarly impaired by visual distractors preceding and immediately following a unisensory localisation target.

Links between fall risk and impaired multisensory integration

Intriguingly, there is growing evidence of an association between multisensory integration performance and falls in older adults. The reasoning behind investigating this link is that successful navigation and balance is reliant upon the efficient integration of many cues (visual, proprioceptive, vestibular, somatosensory, etc.), and impairment in this process may therefore increase the probability of falling. Setti and colleagues (2011) first demonstrated that older participants with a history of falls were far more susceptible to the sound-induced flash illusion. The group have since shown that balance training can mitigate this effect (Merriman, Whyatt, Setti, Craig, and Newell; 2015) and also demonstrated a link in the opposite direction: that postural sway increases in fall-prone adults when they are presented with incongruent multisensory stimuli (Stapleton et al.; 2013). Separately, Mahoney et al. (2014) found that older adults with slower reaction times, who also showed greater multisensory RT enhancement, had significantly worse static balance scores compared to the faster-responding adults who relied less on multisensory stimuli.

As well as the direct implications for understanding and potentially reducing falls in older adults, this connection is potentially of wider interest as it suggests the possibility of some general mechanism or set of supporting functions for multisensory integration that are not yet understood: it is unlikely that a larger temporal binding window for audiovisual stimuli is directly responsible for an increase in fall risk, so these must instead share a common underling cause.

Summary

To summarise, some multisensory integration processes do appear to be changed with age, but this is a highly variable effect. A basic pattern does emerge, however. In the case of multisensory enhancement effects provided by the addition of an entirely congruent stimulus in another modality, healthy older adults benefit at least as much as do younger adults under all but the most challenging circumstances. Evidence for this is provided primarily by studies of reaction time facilitation and speech comprehension. Conversely, when signals are in some way incongruent, age differences sometimes begin to emerge. This seems to depend to some degree on the nature of the incongruence, with temporal misalignment and speech incongruence more consistently affecting older adults' responses than spatial disparity. Evidence for this comes primarily from studies of multisensory illusions or cross-modal distraction. Very few of the demonstrated age differences are unequivocal, however; most paradigms have examples where the effect is absent or even reversed, and tasks that seem to assess similar underlying mechanisms produce divergent results (e.g. double-flash illusion and temporal-order judgement versus simultaneity judgement). Furthermore, there is an underlying trend across multiple paradigms, but particularly for speech-based stimuli, of visual bias in older adults' responses that is rarely directly investigated.

These findings highlight some critical points. First, it is important that we do not consider multisensory integration as a single process that is uniformly affected by ageing, but as a class of functions and behaviours that may be differentially impaired by age-related changes to supporting mechanisms. Second, there is a need for research that helps to elucidate what specific mechanisms underlie the age differences that are seen, including via the application of neuroimaging and established computational models of multisensory perception. Finally, there are areas where research is far more limited, such as multisensory spatial integration, where new research would be most beneficial.

Thesis overview

In the following empirical chapters I aim to address some of these points. Chapter 2 outlines the methodological foundations of the work. Chapter 3 applies a ventriloquist illusion paradigm, alongside a battery of other behavioural tests, to assess older adults' audiovisual spatial integration performance. The BCI model is fitted to participants' responses to assess their optimality and better evaluate any age differences in localisation/common-source judgement responses, and reaction time data are modelled using a Bayesian compatibility bias approach to improve our understanding of reaction time differences. This study was conducted in collaboration with Ulrik Beierholm and Uta Noppeney. I was responsible for study design, data collection, data analysis, and manuscript preparation. UN was involved in study design, data analysis, and manuscript preparation. UB was involved in data analysis and was primarily responsible for the compatibility bias modelling.

In Chapter 4, we use fMRI to investigate the neural mechanisms underlying older adults' responses to ventriloquist stimuli. In addition to standard mass-univariate GLM analyses, we apply multivariate decoding approaches to better evaluate the nature of the information encoded in various brain regions. This study was conducted in collaboration with Susan Francis, Stephen Mayhew, and Uta Noppeney. I was responsible for study design, data collection, data analysis, and manuscript preparation. UN was involved in study design and data analysis. SF and SM were involved in study design, particularly the optimisation of imaging parameters.

For Chapter 5, we turn our attention to audiovisual speech integration. We use behavioural measures and fMRI to investigate the interactions between age, background noise, and audiovisual asynchrony in measures of sentence comprehension. I conducted this research in collaboration with Giulio Degano and Uta Noppeney. I was responsible for study design, data collection, data analysis, and manuscript preparation. UN was involved in study design and data analysis. GD was involved in study design and is collecting EEG data to compliment the fMRI analyses (these data are not reported in this thesis).

CHAPTER 2: METHODS

In the following chapter I introduce the primary research techniques employed in this thesis. First, I discuss our various approaches towards characterising behavioural responses to audiovisual location and sentence stimuli. I then outline two commonly-used computational models of multisensory integration, the second of which is used to characterise behavioural data in Chapter 3. Following this, I include a short discussion of the older adults who took part in the research, comprising descriptions of the participant databases used and information about the specific samples in each empirical chapter. The final part of the chapter is concerned with MRI methodology. I outline the fundamentals of fMRI blood-oxygen-level dependent (BOLD) imaging of the human brain and describe the mass-univariate and multivariate analyses used in chapters four and five. This is followed by a description and evaluation of some non-standard approaches we took in acquiring and analysing fMRI data, including resting-state fluctuation amplitude (RSFA) correction for vascular reactivity, and the temporally-sparse echo-planar imaging (EPI) technique applied in the fifth chapter to allow for the presentation of auditory stimuli in a quiet context during an fMRI experiment.

Characterising behavioural responses to multisensory stimuli

Audiovisual localisation

Chapters 3 and 4 of this thesis depend on the accurate measurement and comparison of unisensory and multisensory localisation responses. We take three different approaches to this: left/right decision, continuous mouse-response localisation, and discrete button-response localisation. Here I will outline the benefits and drawbacks of each approach, discuss analysis of the resulting data, and briefly review previous research that has utilised them.

Left/right location judgements are easy for participants to perform, requiring only a decision between two buttons (thus limiting any working memory and motor factors, and enabling the measurement of informative reaction-times), and are optimised for the fitting of psychometric functions that provide a direct measure of both accuracy and left/right bias. We therefore use this approach to characterise participants' spatial hearing ability. In this procedure, sounds are presented randomly from a number of locations along the azimuth (either via concealed speakers, as in Chapter 3, or via headphones using simulated location cues, as in Chapter 4), and participants instructed to press a button to indicate whether the sound originated from the left or right side of space. These responses are then characterised in terms of the probability of the participant responding that the stimulus is on the right when presented in each location (near zero for extreme left locations; near one for extreme right locations) and a cumulative probability distribution (generally a cumulative Gaussian) is fitted to these probability values. The research presented in this thesis uses the Palamedes toolbox (Prins & Kingdom, 2018), one of a number of tools available for this purpose (e.g. psignifit; Schütt, Harmeling, Macke, & Wichmann, 2016), to fit the distribution using maximum likelihood estimation. This distribution that describes participant responses is known as a psychometric function, and may be formulated as:

$$\psi(x;\alpha,\beta,\gamma,\lambda) = \gamma + (1-\gamma-\lambda)F(x;\alpha,\beta)$$

where, in our case, x refers to sound location, ψ refers to the probability of responding "right", and F is the cumulative Gaussian. γ and λ are nuisance parameters that describe the asymptotes of the function, thus characterising accidental incorrect button presses when the stimulus was on the left or right respectively (i.e. participant pressed "right" when they perceived the stimulus as on the left, and vice-versa). As recommended by Wichmann and Hill (2001), we constrain γ and λ to fall between 0 and 0.05. The remaining parameters are α, which is the displacement of the function to the left or right, and β, which describes the slope. From this fitted distribution we can derive our measures of interest: the point of subjective equality (PSE), which describes left/right bias, is simply α; the just-noticeable difference (JND), a measure of accuracy, is calculated as the difference between the PSE and the value x of the function at a probability of 0.84. The JND is usually specified at this 0.84 level because it corresponds to the standard deviation of the underlying Gaussian distribution (Ernst, 2006). See Figure 3.1A for examples of psychometric function fits to left/right sound localisation responses. We also use this approach to characterise participants' ability to estimate the location of transient visual stimuli by replacing the sound with a cloud of dots sampled from a bivariate Gaussian distribution; the spatial reliability of the stimulus, and thus the JND, may be manipulated by changing the standard deviation of this distribution.

The left/right, forced-choice approach to characterising unimodal localisation ability is a powerful one: it can provide very precise estimates of both discrimination threshold and left/right bias. Psychometric function fits to forced-choice tasks are also frequently used to measure sensory cue weighting in forced-fusion (i.e. stimuli are always perceived as having a single source) multisensory tasks (see e.g. Alais & Burr, 2004; Ernst & Banks, 2002). The method is not well suited to Bayesian Causal Inference modelling, however, which benefits from the richer information provided by approaches that allow for more freedom in spatial responses.

One way to provide this extra information is to use continuous responses, such as in the mouse-cursor method used for both unimodal and bimodal (ventriloquist) localisation tasks in Chapter 3. In this case, unimodal sound localisation performance is measured by requiring participants to move a mouse cursor to indicate the perceived location of a sound presented randomly from one of various locations on the azimuth. To disentangle the various factors that may contribute to inaccuracy in this task we can fit a standard linear regression model to the response data Y:

$$Y = a + bX$$

True sound location is used as the predictor variable *X*. The intercept a of the fitted line provides a measure of left/right bias, while the slope b indicates central bias (the degree to which the participant perceives sounds as being closer to the centre of space than they truly are). The root mean squared error (RMSE) between the true responses and the fitted line is a measure of response variability or noise:

$$RMSE_{fitted} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

Finally, the RMSE between responses and true sound location can be considered a measure of absolute accuracy:

$$RMSE_{true} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}}$$

See Figure 3.1D for an example. Note that this approach provides a new measure, central bias, that cannot be accessed using left/right judgement tasks alone. Deconstructing continuous responses to unimodal stimuli in this way therefore allows us to better describe the specifics of any group differences between younger and older adults. As with left/right responses, this approach may also be used for responses to unimodal visual stimuli of various reliability levels.

When auditory and visual stimuli are presented simultaneously at multiple locations and with varying degrees of spatial disparity, and participants asked to selectively respond to one modality, continuous mouse responses may also be used to measure the influence of one modality on the perceived location of another (ventriloquist effect). We can characterise this as a ratio known as crossmodal bias (CMB), which gives the attraction effect as a proportion of the distance between the two signals. For the most common case of visual stimuli influencing perceived auditory location, this is calculated as:

$$CMB = \frac{A_{responded} - A_{true}}{V_{true} - A_{true}}$$

This value, plotted as a per-condition mean for various spatial disparities and sensory reliability levels, gives a profile of a participant's (or group's) responses to audiovisual spatial stimuli (see Figure 3.3A). BCI modelling, described in a later section, can provide further information about this behaviour.

Mouse-response localisation has some significant drawbacks, however. Responses are much noisier than button presses due to several factors that are impossible to separate. When stimuli are transient, as ours are, the increased time needed to give a mouse response means that working memory plays a much greater role, and motor noise is far more influential when responses are not constrained. The nature of mouse responses also means that artificial trial-to-trial autocorrelation is introduced to the data, as the cursor start location for any response is dictated by the previous trial. Common approaches to combating this can create new problems: randomising the cursor location between trials disturbs the mapping between mouse and cursor, which our piloting demonstrated was particularly disruptive for older adults; requiring participants to return the cursor to the centre between trials may enhance central bias and increase the time needed to run an experiment; and replacing the mouse with a trackball means participants are often using an unfamiliar piece of equipment, which thus increases response times and noise. The longer, more deliberate movements also mean that response time information is usually uninformative, with effects of condition drowned out. Finally, this response type is unsuitable for use in functional imaging experiments due to the long and variable response times, and because the cursor may elicit significant confounding eye movements and visual activations.

The final response type that I will describe is well-suited for use in an imaging experiment, however, and may be considered a compromise between the above two approaches. Discrete button-response localisation, which is used for the MRI paradigm in Chapter 4, has the benefits associated with button responses while also providing some of the more detailed location information necessary for fitting the BCI model and measuring ventriloquist effect at a wide variety of spatial separations. Participants are provided with an array of buttons and told that they map to specific spatial locations; they must press the button that most closely approximates the source location of the sound they perceived. Pilot testing revealed that our participants required some training to learn this mapping, and that some older adults struggle to do so for more than four buttons (this may be an effect of relatively limited experience with computer keyboards/game pads etc. and not age per se). Nonetheless, during the behavioural session described in Chapter 4 participants from both age groups learned to respond with button presses to simulated spatial cues in four locations, each separated by 10°, with high accuracy (above 90%) after only a short period of practice.

Responses may be quantified in various ways when buttons are used to indicate discrete locations. One possibility is raw accuracy: as a percentage, how often did the participant press the button that corresponded to the true stimulus location? This provides limited information about biases etc. but is sufficient to measure the effect of e.g. localisation training (Chapter 4 behavioural session). It is also possible to calculate CMB for discrete button responses. The approach that best suited our requirements, however, was to characterise and plot responses in terms of per-condition mean (see Figure 4.1B for an example). This had the benefit of providing a measure that was directly analogous to the predicted labels provided by our support vector regression fMRI analysis (described later). This form of button response is also commonly used in BCI modelling studies (e.g. Koerding et al., 2007; Rohe & Noppeney, 2015), though in the case of our MRI study applying the model would have been challenging and of limited usefulness due to the symmetrical design.

Audiovisual sentence comprehension

The study described in Chapter 5 uses behavioural testing and fMRI to assess the benefit of multisensory integration for the perception of sentence stimuli under various conditions of background noise and audiovisual asynchrony. A direct and objective approach to quantifying sentence comprehension is for participants to repeat as much of the stimulus as they understood immediately after it is presented, and for this response to be given an accuracy score (see e.g. Grant & Greenberg, 2001; McGettigan et al., 2012; Maguinness et al., 2011). This approach is impractical during an fMRI session, however, so it was necessary for us to use a subjective rating system to allow participants to indicate their comprehension of audiovisual sentence recordings while being scanned. In order to verify that participants' ratings matched their true level of comprehension we developed and administered a novel paradigm that quantified the correspondence between subjective and objective responses. During the behavioural screening/testing session that preceded the MRI scan, participants took part in a sentence comprehension task that required dual responses. Following the presentation of a sentence in one of a number of conditions (unimodal auditory, unimodal visual, synchronous bimodal, or asynchronous bimodal, with or without background noise; see Chapter 5 for details of the stimuli and conditions), participants were instructed to: (a) press one of three buttons to indicate their comprehension: *understood none*, *understood some*, or *understood fully*, then (b) to repeat the sentence as accurately as possible to the experimenter. This verbal response was marked on a sheet generated automatically for each participant, assigning a score based on the number of target words (out of a possible total of five per sentence) that were correctly reported. To assess the correspondence between the two responses, we entered each set of scores into separate correlational analyses for noise and no-noise conditions. Though we set this correspondence score as an exclusion criterion, all participants tested achieved scores above *r* = 0.7 in either noise condition, confirming that our in-scanner subjective-report paradigm is a good index of true sentence comprehension. See Figure 2.1 for an illustration of this between-task correspondence.

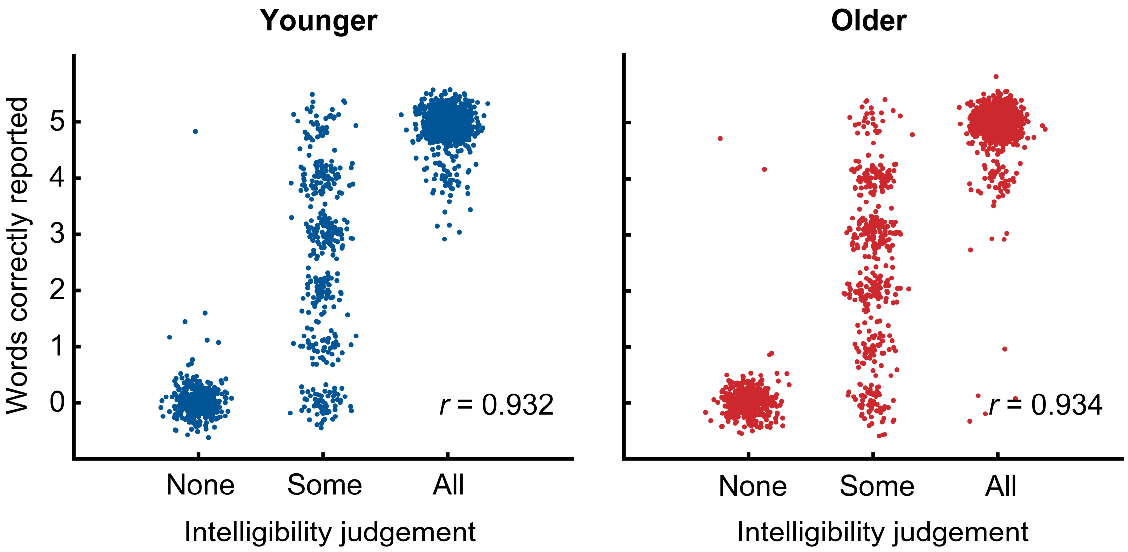

Figure 2.1. Correspondence between sentence repetition and comprehension report tasks. Plotted data points are individual trials from all subjects, randomly jittered for illustration. Correlation coefficients are based on single responses concatenated across participants.

Computational modelling of multisensory integration

Maximum likelihood estimation

In the past two decades, considerable advances have been made in the mathematical modelling of multisensory perception. Particularly, it has been demonstrated that human behavioural responses to multisensory stimuli may be closely approximated by statistically optimal computational models. Building on earlier work on the integration of multiple visual intramodal cues (e.g. Jacobs, 1999), Ernst and Banks (2002) demonstrated that visual-haptic height estimation followed the predictions of a maximum-likelihood estimator (MLE). Alais and Burr (2004) later found the same for audiovisual spatial integration.

This *forced-fusion* model, so called because it focuses on cases in which two cues are integrated (not processed separately), assumes that the information in each sensory modality i about some physical property is based on an estimate $\hat{S}_i$ that is corrupted by Gaussian noise, and that the variance $\sigma^2{}_i$ (i.e. uncertainty) associated with each sensory modality's estimate may be different. A mathematically optimal observer would construct the final joint estimate as an average of the two estimates, weighted by their respective reliabilities. For one auditory and one visual stimulus,

$$\hat{S}_{AV} = w_A \hat{S}_A + w_V \hat{S}_V$$

where

$$w_A = \frac{1/\sigma_A^2}{1/\sigma_A^2 + 1/\sigma_V^2} = \frac{\sigma_V^2}{\sigma_A^2 + \sigma_V^2}$$

and vice-versa for w_V. A major prediction of this model is that the variance of the combined estimate will always be smaller than either of the unisensory estimates,

$$\sigma_{AV}^2 = \frac{\sigma_A^2 \sigma_V^2}{\sigma_A^2 + \sigma_V^2} < min(\sigma_A^2, \sigma_V^2)$$

representing multisensory enhancement of the estimation of the physical property.

The experimental approach towards assessing whether participants' responses align with the predictions of this model generally involves collecting responses to unisensory stimuli in both modalities and using the above equations to calculate MLE-optimal joint estimates and variances. These are then checked against real responses to bimodal stimuli, presented with an imperceptible level of conflict to measure how much weight participants are placing on each sensory modality in their final estimate (Alais & Burr, 2004; Ernst & Banks, 2002; Helbig & Ernst, 2008; Jacobs, 1999; Takahashi, Diedrichsen, & Watt, 2009).

Bayesian causal inference

Generative model and inference

While an excellent way of assessing optimality of multisensory integration in a forced-fusion context, this model and paradigm is somewhat limited by the fact that it cannot account for situations in which sensory cues should not be integrated but processed separately (segregated), which in the real world is true for the majority of incoming signals at any one time. With this limitation in mind, Koerding et al. (2007) developed a model that was based on the established reliability-weighted integration of signals, but also allowed for the possibility that multiple signals originate from separate sources. Though it has a general form, this Bayesian causal inference (BCI) model has been applied in various ways depending on the type of data and responses collected (Elliott, Wing, & Welchman, 2014; Peters, Ma, & Shams, 2016; Shams, Ma, & Beierholm, 2005); here I will outline the approach used in Chapter 3, where we fit the BCI model jointly to continuous (mouse cursor) auditory localisation and common-source judgements. The stimuli are one visual and one auditory signal per trial, presented at multiple levels of azimuthal disparity, with varying visual reliability (see Chapter 3 methods section for further details of the paradigm).

The BCI model creates predicted responses based upon a model of the underlying structure of the world and sensory system. It is optimised to a given behavioural dataset by adjusting parameters until the generated predictions best match the behavioural responses. For a pair of signal estimates, in our case a transient sound and flash of dots, there are two possible causal structures: either they had the same source, and should therefore be integrated, or were produced by separate sources and should be segregated. The BCI generative model explicitly models both possibilities: it determines whether signals have common ($C = 1$) or separate ($C = 2$) sources by drawing from a binomial distribution determined by a common-source prior, $P(C = 1) = p_{\text{common}}$. For signals with a common source, their shared location S_{AV} is sampled from prior distribution $N(\mu_P, \sigma^2{}_P)$. We assume a perceptual bias towards the centre of space, so $\mu_P = 0$. For signals with independent sources, S_A and S_V are drawn independently from this prior distribution. Sensory noise is then introduced by sampling spatial estimates x_A and x_V from normal distributions centred respectively at S_A and S_V, with variances $\sigma^2{}_A$ and $\sigma^2{}_V$. This generative model therefore has four free parameters: the prior probability of a common source p_{common}, the variance of the spatial prior $\sigma^2{}_P$, the variance of the auditory signal $\sigma^2{}_A$, and the variance of the visual signal $\sigma^2{}_V$. If multiple levels of visual reliability are present, as in our case, the final parameter is fitted separately for each.

Bayes' rule can then be used to infer the posterior probability of the signals' causal structure (Koerding et al., 2007):

$$p(C = 1|x_A, x_V) = \frac{p(x_A, x_V|C = 1)p_{common}}{p(x_A, x_V)}$$

When $C = 1$ (the signals share a source), the optimal estimate of the sound's location is the reliability-weighted average of the auditory estimate, the visual estimate, and the spatial prior:

$$\hat{S}_{AV,C=1} = \frac{x_A/\sigma^2{}_A + x_V/\sigma^2{}_V + \mu_p/\sigma^2{}_P}{1/\sigma^2{}_A + 1/\sigma^2{}_V + 1/\sigma^2{}_P}$$

When $C = 2$, the auditory location estimate is independent of the visual signal:

$$\hat{S}_{A,C=2} = \frac{x_A/\sigma^2{}_A + \mu_p/\sigma^2{}_P}{1/\sigma^2{}_A + 1/\sigma^2{}_P}$$

Decision strategies

To translate these two posterior probabilities into final estimates of a sound's location, the brain must combine or differentiate between them in some way. Wozny et al. (2010) describe three possible decision functions or strategies: *model averaging*, *model selection*, and *probability matching*. The model averaging strategy assumes the brain combines the integrated ($C = 1$) and segregated ($C = 2$) signal estimates weighted by the posterior probability of their underlying causal structure:

$$\hat{S}_A = p(C = 1|x_A, x_V)\hat{S}_{AV,C=1} + (1 - p(C = 1|x_A, x_V))\hat{S}_{A,C=2}$$

Model selection instead assumes that the observer simply selects the estimate from the most probable causal structure:

$$\hat{S}_A = \begin{cases} \hat{S}_{AV,C=1} \; if \; p(C = 1|x_A, x_V) > 0.5 \\ \hat{S}_{A,C=2} \; if \; p(C = 1|x_A, x_V) \leq 0.5 \end{cases}$$

Finally, the probability matching strategy assumes that the brain selects between the two estimates probabilistically based on the relative probabilities of the two causal structures. This is modelled by randomly drawing a selection criterion ξ from a uniform distribution between 0 and 1 on every trial.

$$\hat{S}_A = \begin{cases} \hat{S}_{AV,C=1} \; if \; p(C = 1|x_A, x_V) > \xi \\ \hat{S}_{A,C=2} \; if \; p(C = 1|x_A, x_V) \leq \xi \end{cases}$$

In Chapter 3 we fitted the BCI model using each of these decision strategies. As model averaging was the winning strategy across participants in both age groups (based on group-level protected exceedance probabilities; Rigoux et al., 2014), we report results based on this strategy. For the common-source responses, we assumed participants would report that the auditory and visual signals shared a source whenever the posterior probability of a common source exceeded 0.5:

$$\hat{C} = 1 \; if \; p(C = 1|x_A, x_V) > 0.5$$

$$\hat{C} = 2 \; if \; p(C = 1|x_A, x_V) \leq 0.5$$

Optimising the parameters

Predicted distributions of common-source and auditory location estimates are produced by marginalising over x_A and x_V. These distributions are generated by simulating x_A

and x_V many times (in our case 10,000 iterations) per condition and calculating $\hat{C}$ and $\hat{S}_A$ using the above equations. The distribution of predicted auditory locations can then be further perturbed to simulate presence of motor noise in the continuous spatial responses; we used a Gaussian kernel with width $\sigma = 1°$ (in Chapter 3, fitting the model with this as an extra free parameter led to some changes in parameter estimates but did not introduce any between-group differences). The log-likelihood of participants' sound localisation and common-source responses given the model may then be computed and summed over conditions.

A two-stage fitting process is used to optimise the model parameters to each participant's responses (i.e. maximise the log-likelihood). First, a grid search is run across various likely settings for each of the parameters. The parameters with the highest log-likelihood are then used as a starting point in an optimisation algorithm (we used fminsearch as implemented in MATLAB R2014b) that attempts to find the set of parameters that produces simulated data which most closely resembles the participant's true responses. To investigate age differences in audiovisual spatial localisation behaviour we compared these fitted parameters between groups.

Conclusion

The BCI model is a powerful approach that has been used to model many types of data (see Chapter 1 for examples). It acknowledges and accounts for the fact that most signals should be processed separately while still allowing for forced-fusion cases in which stimuli are fully integrated, and explains the entire spectrum of partially integrated percepts that can occur in between. It is by no means perfect, however. Wozny et al.'s (2010) discovery that participants seem to recruit different response strategies suggests there are top-down (e.g decisional, attentional) influences on participant responses that the model cannot account for. Furthermore, several studies (e.g. Burr et al., 2009; Butler et al., 2010; Rosas et al., 2005) have found sub-optimal weighting of signals in a forced-fusion context, suggesting the reliability-weighted integration on which the model relies may not be entirely robust. It is nonetheless the best-performing general model of multisensory perception that currently exists and fits the data in Chapter 3 of this thesis very well (see e.g. Figure 3.3).

Recruiting older adults

The findings of cognitive ageing research will depend to a degree on the properties of the population sampled. Older adults living in residential accommodation are, for example, more likely to exhibit cognitive impairment and other psychiatric illness compared to those living in the community (Mann, Graham, & Ashby, 1984; Nikmat, Al-Mashoor, & Hashim, 2015; Sutcliffe et al., 2007), and previous occupation has been shown to be a predictor of cognitive impairment (Jorm et al., 1998). In this section I therefore outline the specific participant databases used for recruiting participants to the empirical research described in this thesis, and explain how these databases were utilised to various degrees for each study.

Participant databases

Lifespan Cognition Database

The University of Birmingham School of Psychology maintains a database of older adults in the local area who are willing to take part in psychology/neuroscience research. This Lifespan Cognition Database is actively maintained by a participant coordinator who advertises both within and beyond the university, attending groups such as the local ramblers' society and Women's Institute, to recruit new members. Many are also recruited through word-of-mouth via existing members. This varied approach results in considerable diversity in the older adults' backgrounds, qualifications, and experience. Upon signing up, all members routinely complete a questionnaire providing information about their medical history, current medication, level of education, and previous occupations, allowing researchers to specify some inclusion criteria (e.g. right-handed, no history of psychiatric disorders) before participants are approached to take part in a study. The research described in Chapter 3 was conducted on older participants recruited exclusively from this database.

1000 Elders

A larger database of participants, in existence since the early 1980s, is also maintained by the university's College of Medical and Dental Sciences. Methods of recruitment are comparable to those used for the Lifespan Cognition Database (though the 1000 Elders is also advertised online and at annual public engagements events at the university), so the members are similarly diverse. The information stored about 1000 Elders members is more limited, and studies advertised more widely (i.e. without any pre-selection), so the number of reported exclusions is likely to be higher. Many of the older participants who took part in the research described in Chapters 4 and 5 were recruited via this database.

Sample characteristics for each study

As indicated above, all older participants who took part in the research for Chapter 3 were recruited via the Lifespan Cognition Database. However, it was necessary to also advertise to the larger 1000 Elders database for the study reported in Chapter 4, as many older adults are not willing or able to take part in extended sessions of fMRI scanning. The more general approach to advertising research used by this database, combined with our stringent eye-tracking criteria, led to a relatively high number of exclusions in this study. Thus, having already screened and identified suitable participants, we elected to re-recruit as many as possible to the second fMRI study (reported in Chapter 5). A small number of participants were also recruited via other means, such as existing participants advertising the research to friends and family. See Table 2.1 (below) for more details.

Table 2.1

Details of older participant recruitment for each study

	Chapter 3	Chapter 4	Chapter 5
Total participants recruited	24	29	15
Lifespan Cognition Database[1]	24	7	4
1000 Elders[1]	0	27	14
Word-of-mouth/other	0	2	1
Participants included (i.e. after exclusions)	20	16	15
Shared with Chapter 3	-	0	1
Shared with Chapter 4	0	-	12
Shared with Chapter 5	1	12	-

[1] *Some participants are members of both databases*

Magnetic Resonance Imaging

Functional blood-oxygenation-level dependent imaging

Developments in MRI technology in the past decades allow us to measure changes of dynamic blood flow in the brain to estimate neural activity (this is referred to as functional MRI, or fMRI). When a group of neurons becomes active, the amount of oxygen-rich blood flowing to the region is increased. Usefully, the magnetic properties of blood vary depending on the amount of oxygen it is carrying. By rapidly and repeatedly acquiring images of the brain using MRI sequences sensitive to these magnetic properties, we can measure changes in blood oxygen levels over time. This is referred to as blood-oxygenation-level-dependent (BOLD) imaging. It is important to note that the physiological processes underlying this response are still not fully understood, and BOLD signal changes should be considered only an indirect measure of neural activity (Logothetis, 2008): it can be reasonably claimed that a change in BOLD signal correlates with neural activity, but the precise meaning of, for example, a negative BOLD signal can be difficult to interpret (Mullinger, Mayhew, Bagshaw, Bowtell, & Francis, 2014).

A stimulus-evoked BOLD response does often adhere to a general shape, however (Lindquist, Meng Loh, Atlas, & Wager, 2008). Following a small initial dip below baseline in the first 1-3 seconds after stimulus onset, the signal change increases and peaks at somewhere near five seconds, before decreasing at a slightly slower rate, then undershooting below, and finally returning to, baseline. This pattern of activity is known as the haemodynamic response function (HRF). The primary feature of note is the slowness of the response (20 or more seconds between onset and return to baseline) in comparison to the neural activity that evoked it, which may have lasted only milliseconds. This, along with physical constraints on the time taken to acquire images, necessarily means that the temporal resolution of fMRI is on the order of seconds, as opposed to the potentially sub-millisecond accuracy of electrophysiological imaging methods such as EEG, MEG, or ECoG. However, assuming stimulus onsets are sufficiently spaced in time, the haemodynamic response function may be considered (and modelled) as a linear time-invariant system (Poldrack, Mumford, & Nichols, 2011); the shape of the response is independent of onset time, and its magnitude is both additive and scales linearly with the size of the neural response. In practice, this means that it is possible to model the BOLD response to multiple events that occur within the duration of a single HRF. The specifics of this modelling approach are discussed in the later section on the general linear model.

fMRI data analysis

A number of software packages and toolboxes have been used throughout this thesis for the preparation and analysis of brain imaging data. SPM12 (Friston et al., 1994) was used for pre-processing and general linear modelling. The Decoding Toolbox (TDT; Hebart, Görgen, & Haynes, 2015) was used for multivariate analysis. FreeSurfer (Fischl, 2012) was used for the creation of subject-specific, anatomically-defined regions of interest (ROIs). These were supplemented extensively by custom MATLAB code.

Pre-processing

The raw functional image files produced by the MRI scanner console must undergo various pre-processing operations before they are suitable for analysis. The operations used, and their parameters, are dependent on both the specifics of the MRI sequence and on the intended analyses. Here I give a summary of the most common preprocessing steps for fMRI date: their purpose, usage cases, and a brief description of the operations involved.

Spatial realignment and unwarping

Functional imaging data consists of a series of images acquired over a period of time, the changes between which are used to estimate activity. However, and despite the use of physical measures such as padding to control movement, human subjects are not often perfectly stationary throughout a scanning session. This means that any signal change at a specific point in space may simply be due to movement in the tissue being measured. It is also often desirable to make inferences about activity across multiple scanning sessions, or even multiple days, the images from which will certainly not be in alignment. The first stage in most functional preprocessing pipelines is therefore the spatial realignment of all volumes acquired, to ensure that each voxel (volume unit; a three-dimensional analogue of a pixel) contains the same area of tissue. This is a mathematically straightforward operation: each image in the series undergoes rigid body transformations until the mean squared difference between it and a reference image (usually the first in the series) is minimised. The parameters of this realignment may be recorded for later inclusion in the subject-level general linear model (GLM).

Echo-planar imaging (EPI), the MRI sequence type most commonly employed for BOLD fMRI, also suffers from signal dropout and distortion artefacts around the areas of the brain that are located near to air-filled cavities, such as the more anterior and inferior parts of the prefrontal cortex (PFC). These distortions interact in a non-linear way with participant movement. SPM12 therefore includes an algorithm that attempts to account for and correct these “susceptibility-by-movement” interactions by modelling them and applying the inverse of the calculated deformation fields to the images (Andersson et al., 2000).

Slice-timing correction

The acquisition of a full functional brain volume generally takes between 2 and 3 seconds, depending on the scanner and sequence parameters. Images consist of a series of slices, acquired in sequence, meaning that the last slice is acquired seconds later than the first. For example: if the total image acquisition time (TA) for a brain volume is 2.8s and the slices are acquired in ascending order (i.e. inferior to superior), the BOLD signal at the primary motor cortex is measured almost three seconds after that at the cerebellum. It is therefore often useful to apply a temporal interpolation such that the measured signal in each slice is ‘shifted’ in time relative to a reference slice (usually the middle), resulting an image that simulates simultaneous slice acquisition. This adjustment is applied uniformly to all volumes in a series.

Tissue segmentation/spatial normalisation

Though arguably distinct operations, I have grouped tissue segmentation and spatial normalisation under one heading as SPM12’s standard algorithm takes a unified approach to these operations, performing them simultaneously in an iterative fashion. Before they are applied, the functional and anatomical images are coregistered to the same space by applying rigid-body transformations until their mutual information is maximised (Ashburner & Friston, 2005). This allows the below steps to be performed using each participant’s highly-detailed structural image, and the resulting transformations to be applied to the lower-resolution functional data, greatly increasing accuracy.

The purpose of SPM12’s segmentation/normalisation algorithm is to correct intensity biases across a brain image, to separate the image into various constituent tissue classes, and to deform (‘normalise’) a participant’s brain image to match a standard brain image; the commonly-used Montreal Neurological Institute (MNI) space has been used for spatial normalisation throughout this thesis. These processes are conducted within the same algorithm because their dependencies are circular: tissue segmentation relies on spatial normalisation, but accurate normalisation requires segmented tissues, et cetera.

Tissue segmentation is the identification and separation of the various tissue types (grey matter, white matter, cerebrospinal fluid (CSF), bone, etc.) in an image. As well as aiding with spatial normalisation, these segmented images are useful for various subsequent analyses (including the resting-state fluctuation correction discussed later in the chapter). Segmentation is accomplished by the calculation of tissue interface boundaries based on various overlaid tissue probability maps. Spatial normalisation is of particular importance to group-level analysis, as it allows for a group of subjects with considerable differences in brain morphology to be analysed (and reported) in the same way. It is accomplished through a non-linear transformation of the participant’s anatomical image to match a target (MNI) volume (Ashburner & Friston, 2005).

Spatial smoothing

The final step of many preprocessing pipelines is the spatial smoothing of functional images. Each volume is convolved with a three-dimensional Gaussian kernel of a given width. This serves both to improve the signal-to-noise ratio (SNR) of images and, in the case of group-level analyses, to further improve spatial correspondence between participants. The degree of spatial smoothing used is dependent on the intended analysis. Here we have used an 8mm kernel for group-level analyses, to maximally reduce the impact of residual functional differences between participants after MNI normalisation. Conversely, we have used only 3mm smoothing when conducting multivariate analyses in participants' native anatomical space in order to improve SNR while preserving the important signal differences between voxels.

Mass-univariate GLM analysis

Modelling and statistical analysis of functional images is accomplished by fitting a general linear model (GLM) to the time course data of each voxel individually. Effects of interest such as stimulus onset times are included as regressors in this model, as are nuisance variables such as motion parameters. Regressors that are intended to model neural activity in some way are convolved with a canonical HRF to more accurately model the expected BOLD response. It is also common to include regressors representing multivariate Taylor expansions of the canonical HRF in time (known as the temporal derivative) and width (dispersion derivative) to account for small variations in the latency and duration of the response (Friston et al., 1998). Results of the first fMRI study reported in this thesis (Chapter 4) are based upon GLMs modelling event onsets as a canonical HRF plus its temporal derivative. The second fMRI study (Chapter 5) used only the canonical HRF, as the long repetition time (TR) of the sparse sampling sequence (described in detail later in the chapter) reduces temporal resolution to the point where these derivates are no longer useful.

In fMRI studies that use brief stimuli in an event-related design, such as both of those in this thesis, stimulus onsets may be modelled in separate regressors for each condition of the design. When fitted to the time series data for each voxel, the GLM therefore produces a set of parameter estimates β representing the magnitude of each condition's contribution to the signal. Specific experimental questions may then be addressed by testing contrasts comparing the relative contribution of multiple regressors. As the GLM is run individually for each voxel in the brain, the resulting outputs in fact form a series of three-dimensional maps, including one each for every regressor and contrast.

Random-effects group-level inference is based upon the outputs of these subject-level GLMs. Contrast or beta images for each participant are entered as the data in a new GLM, and further contrasts and statistical tests are conducted on these data; the resulting statistical maps reveal the consistency of results across participants for each contrast of interest. Finally, these maps often undergo multiple comparison correction to account for the statistically probable large number of false positives. SPM12 implements a relatively conservative familywise error (FWE) correction procedure based upon Gaussian random field theory (Worsley et al., 1996), and all second-level results reported in this thesis are at an FWE-corrected level of $p < .05$ unless stated otherwise.

Multivariate analysis

Though the above GLM procedure has proved infinitely useful across decades of brain imaging research, its ability to make inferences based on groups of voxels is necessarily limited by the fact that the model is fitted entirely separately for each voxel. Researchers have therefore moved towards supplementing these traditional univariate methods with more sensitive multivariate approaches that consider the extra information held by collections of voxels activating together in response to a stimulus.

Support vector regression

Because this multivariate pattern analysis (MVPA) involves data with high dimensionality, it often relies on machine learning algorithms, which are well-suited to extracting patterns from large and complex datasets. In Chapter 4 of this thesis we use a regression approach based on support vector machines (SVM). In brief, an SVM considers each piece of data as a point in a multidimensional coordinate system and attempts to find the hyperplane that best separates the data into classes (Smola & Schölkopf, 2004).

To assess the ability of an SVM (or indeed any machine learning algorithm) to correctly classify or predict data, a standard approach is to *train* the machine on one dataset and then *test* the fitted parameters on a separate dataset containing data from the same conditions. This is known as cross-validation. We use *k*-fold leave-one-run-out cross validation (as implemented in The Decoding Toolbox), where *k* is equal to the number of runs. For example, if an fMRI task has been split evenly into eleven separate runs (as in the study in Chapter 4), the following training/testing scheme is used. A support vector machine is trained on labelled imaging data from the first ten runs of the task, then tested on imaging data from the eleventh for accuracy. This is repeated eleven times, leaving a different run out for testing each time, and the resulting eleven accuracy measures averaged to produce an overall accuracy value for the model.

The specific purpose of support vector regression (SVR) as opposed to classification approaches is to predict the value of a data point on a continuous scale. This is very useful for localisation tasks that have stimuli appearing at multiple possible locations. For example, an SVR model trained on BOLD responses to auditory stimuli in a variety of locations that is then tested on new data will predict the exact location of the stimulus that produced the new response (unconstrained by pre-defined categories). The accuracy of this prediction is of course dependent on the quality of the data: BOLD response data from primary auditory cortex or planum temporale will produce more accurate predicted labels for auditory location than BOLD data from primary visual cortex. Similarly, an age difference in the accuracy or reliability of these predicted labels might suggest that older adults have a noisier representation of auditory location in the areas of interest.

Multivariate Bayes

Many fMRI studies of ageing report group differences in BOLD activation (see Chapter 1 for examples). Greater BOLD activity in older adults is generally interpreted as being in some way compensatory. It is increasingly acknowledged, however, that this is not the only interpretation: it may also reflect simple activation inefficiency (Morcom & Johnson, 2015). In order to differentiate between these possibilities, it is necessary to investigate whether extra regions being recruited by older adults encode more information about the stimuli or task compared to (a) the same regions in younger adults and (b) other regions activated similarly by both age groups. That is, if the regions with greater activation in older adults actually hold similar amounts of relevant information in both age groups, or encode no extra information compared with regions that both groups recruit, the activity increase cannot be considered compensatory. Support vector machines are not well suited to addressing this problem as they cannot be characterised in terms of model evidence, precluding formal model comparison (Friston et al., 2008; Morcom & Johnson, 2015). The Multivariate Bayesian (MVB) decoding approach described by Friston et al. (2008) and implemented in SPM12 does allow for this. MVB uses the same data Y and design matrix X used for mass-univariate GLM analysis but inverts the direction of prediction. Instead of regressors in X attempting to predict activations in Y, the data in Y attempt to predict an outcome, or target, variable contained within X. This target variable is specified as a standard contrast; all other modelled regressors are removed as nuisance variables. A hierarchical parametric empirical Bayes model then attempts to find the pattern of voxel weights that best maps the BOLD data across many voxels to the target variable while also optimising the set size in terms of free energy.

The outcome variable is log model evidence, which can be used to directly compare the amount of information about the contrast of interest held in various regions. See Friston et al. (2008) for more details, and Morcom et al. (2018) for an example of implementation in ageing research.

Sparse sampling

The standard approach to task-related fMRI is to acquire brain volumes constantly throughout a run while stimuli are presented at defined intervals. Each volume is acquired immediately following the previous or, put another way, the acquisition time (time to acquire a whole volume; TA) and the repetition time (time between starting acquisition of each volume; TR) are near-identical. As long as these stimulus onsets are jittered, or have a stimulus-onset asynchrony (SOA) that is different from (and not a multiple of) the TR, this continuous scanning allows for accurate characterisation of the BOLD response to stimuli by sampling it at multiple points. This approach has some disadvantages, however. Of particular interest here is the loudness of the scanner: the continuous switching of gradients necessary to acquire fMRI images is extremely noisy. The MR-safe headphones used to present sound stimuli inside the scanner usually feature sufficient passive noise isolation to reduce sound pressure at the ear to safe levels, and the multisensory localisation data in Chapter 4 were collected under these conditions; both younger and older adults were still able to accurately localise sounds under these relatively loud conditions. (Many MRI studies employ noise-attenuating earplugs to further reduce sound pressure at the ear, but these can interfere considerably with sound localisation; see e.g. Zimpfer & Sarafian, 2014.)

However, piloting revealed that scanner noise was a significant confounding factor in the multisensory speech task described in Chapter 5. Even with extensive hearing protection, scanner noise was found to significantly impair speech comprehension for participants in the ostensibly noise-free conditions. This was a particular problem for a study which aimed to compare speech comprehension in noisy versus noise-free conditions, especially as it disproportionately affected older adults. We therefore opted to use a temporally-sparse EPI sequence that featured a long 8800ms TR with 2800ms of slice acquisitions clustered at the beginning (Edmister et al., 1999; Hall et al., 1999). These parameters resulted in a 6000ms between every volume acquisition, during which stimuli could be presented in near-silent conditions.

Sparse sequences have several disadvantages when compared to standard continuous EPI (Peelle, 2014). The most obvious is the simple loss of power due to a lower sampling rate: a continuous version of the above sequence that acquires a new volume every 2800ms would collect over three times as many images during the same time period. We mitigated this issue by increasing the number of stimuli per condition, therefore extending the duration of each scanning session. Sparse sampling also results in a reduced ability to characterise the shape of the BOLD response to a stimulus: using our sparse sequence described above, the response to any one stimulus is captured at only a single time point (with perhaps a small amount of residual activation apparent in the following volumes). This could prove problematic in ageing research, as there is some evidence of age differences in the shape and timing of the HRF (Aizenstein et al., 2004; Ances et al., 2009; Ward, Aitchison, Tawse, Simmers, & Shahani, 2015, though also see D'Esposito, Zarahn, Aguirre, & Rypma, 1999); when sampled at only a single time point, a slower BOLD response could be misinterpreted as reduced activity. To reduce this problem, our sequence purposefully included a period of silence (6000 ms) that was longer than the duration of the stimuli (2197 – 3804 ms). This allowed the onsets of the stimuli to be jittered, and therefore the BOLD response to be sampled at different points, resulting in a better characterisation of responses across the multiple stimuli within a condition.

Resting-state fluctuation amplitude correction

One of the greatest challenges when using BOLD fMRI to measure differences in neural activation between younger and older adults is age-related variations in vasculature and neurovascular coupling. As described earlier, the BOLD signal measures neural activity indirectly via variations in blood oxygenation, making it sensitive to individual differences in cerebral vasculature and vascular reactivity. Each of these has been shown to vary systematically with age (Riecker et al., 2003). The practical implication of this is that any observed group differences in raw BOLD activity may conceivably be due entirely to these vascular factors and not neural activity. It is therefore necessary to apply some form of correction to mitigate the impact of any age differences in blood flow. The most common approach to correcting for individual variation in cerebral blood flow is to divide the BOLD response by a hypercapnic baseline, induced by either an extended breath hold or by CO_2 inhalation (Liu, Glover, Mueller, Greve, & Brown, 2013). Each of these approaches has significant disadvantages, however. Gas inhalation is time-consuming and requires specialist equipment, while breath-hold challenges are subject to considerable interindividual variability, particularly in older adults (Riecker et al., 2003). Both also involve a task, which will induce some neural activity (Hall et al., 2011), and neither are especially well-tolerated by older populations (Liu, Glover, Mueller, Greve, & Brown, 2012; Tsvetanov et al., 2015). We instead elected to use resting-state fluctuation amplitude (RSFA), a measure first described by Kannurpatti and Biswal (2008) and later validated by Tsvetanov et al., (2015), to correct our BOLD fMRI images for individual differences in cerebral vascular reactivity. RSFA is easier for older participants and far less time-consuming, as it only requires the acquisition of a short (5 min) resting-state scan. Nonetheless, Kannurpatti and Biswal found the technique had over 80% correspondence with response to hypercapnic challenge (based on both within and between-participant variance).

The basic idea of RSFA correction is to use variability in resting-state BOLD signal to calibrate task-induced responses. Tsvetanov and colleagues describe a pipeline for the preparation of RSFA maps. Resting-state data are collected during the same scanning session, using the same sequence parameters, as the task data. Images are initially preprocessed in the same way as the task data, including spatial realignment, slice-timing correction, segmentation/normalisation, and spatial smoothing. Extra steps are then applied to further minimise non-vascular contributions to the signal. The authors describe wavelet despiking (Patel et al., 2014), regression of movement parameters (and their first derivative), regression of mean white matter and cerebrospinal fluid signal (based upon SPM12 probability maps thresholded at 0.75), and linear and quadratic detrending. A bandpass filter of 0.01 – 0.08Hz is then applied to the time-series data to maximise the contribution of vascular reactivity to the signal. Finally, the standard deviation of the resulting signal at each voxel is calculated to create the RSFA map. Each participant's map is used to scale the parameter estimates (beta images) for each regressor in the first-level GLM.

The RSFA correction applied in Chapters 5 and 6 closely followed this procedure. Resting-state data in Chapter 6 were collected using the same sequence parameters as those in Chapter 5, as they matched the temporally-sparse sequence geometrically but were better suited for measuring physiologically-induced signal fluctuations. Comparisons of GLM analyses conducted with and without RSFA corrections revealed that the effects of this correction on results were moderate: between-group differences were similarly apparent in both cases, but the extent and loci of the effects sometimes varied.

Defining regions of interest

Several of the analyses in this thesis rely on the investigation of BOLD activity in specific regions of interest (ROIs) within the cortex. ROIs may be used, for example, to define boundaries for data to be included in decoding models (Chapter 4), or to restrict mass-unimodal analyses to specific regions (Chapter 5). They may be specified based on anatomical features or on functional activity, and we have employed various versions of each approach depending on the nature region in question and requirements of the analysis. For primary auditory cortex we use maximum probability maps based on cytoarchitectonic (cellular composition) studies of ten adult postmortem brains (Morosan et al., 2001) included in the SPM Anatomy Toolbox (Eickhoff et al., 2005). For visual areas (including V1-V3 and intraparietal sulcus) we also use maximum probability maps, but these are based on retinotopic mapping (functional response patterns to various complex visual stimuli) of 53 healthy adults (Wang, Mruczek, Arcaro, & Kastner, 2015). These maps are available in standardised MNI space; to apply them to participants' native space, they can be inverse-normalised using a deformation matrix that is produced during SPM's unified segmentation-normalisation procedure. The maps are by definition probabilistic, and thus do not provide perfect delineation of the regions of interest, but are more standardised and far less time consuming than alternatives (e.g. manual delineation of regions or per-subject functional localisers). For cortical regions that are well-specified in terms of gross anatomy but variable between participants, and cannot easily be defined with a functional localiser, the surface-based automated cortical parcellation implemented in Freesurfer (Fischl, 2012) is a reliable method of defining ROIs based on gyral and sulcal structure. We use this to define planum temporale in Chapter 4.

Finally, it is often desirable to define functional masks based on contrasts and activations that are study-specific. In this case, maps of significant activations from a specific group-level GLM contrast are thresholded at a defined level and binarised to create a mask. These may then be used in their original form to restrict group-level analyses in the same standard (MNI) space, or inverse-normalised to participants' native space for e.g. multivariate decoding analysis.

CHAPTER 3: OLDER ADULTS SACRIFICE RESPONSE SPEED TO PRESERVE MULTISENSORY INTEGRATION

Samuel A. Jones, Ulrik Beierholm, and Uta Noppeney

Throughout our lifetime we are continually exposed to a barrage of sensory signals. Our ability to navigate through and interact with the world relies on the correct interpretation of this multisensory information. To quickly orient towards and identify a threat, for instance, we may rely on both the sound it makes and its appearance: what creature made that growl, where is it, and is it moving towards me? Previous research has demonstrated differences between younger and older adults in their responses to multisensory stimuli. For example, older adults have been shown to be more susceptible to some multisensory perceptual illusions including the sound-induced flash illusion (DeLoss, Pierce, & Anderson, 2013; McGovern, Roudaia, Stapleton, McGinnity, & Newell, 2014; Setti, Burke, Kenny, & Newell, 2011) and the McGurk-MacDonald effect (Sekiyama, Soshi, & Sakamoto, 2014; Setti, Burke, Kenny, & Newell, 2013), but are less affected by the stream/bounce illusion than younger adults (Roudaia, Sekuler, Bennett, & Sekuler, 2013). The mechanisms underlying these differences are unclear, but it is possible that they are driven in part by age-related changes in sensory reliability (Lindenberger & Baltes, 1994; Liu & Yan, 2007; Salthouse, Hancock, Meinz, & Hambrick, 1996). Race-model analyses of reaction times have also demonstrated that older adults may benefit more from congruent, redundant multisensory stimuli (Laurienti, Burdette, Maldjian, & Wallace, 2006; Mahoney, Li, Oh-Park, Verghese, & Holtzer, 2011), though the relative contributions of changes to integration behaviour, motor speed, and response threshold to these differences have not been established.

Though important, none of the above findings reveal how ageing affects the computations underlying multisensory integration. There are two computational problems the brain must solve when integrating multiple sensory signals. First, it must decide whether the signals are produced by the same object (i.e. share a common source) and should therefore be integrated, or originate from different sources and should be processed separately (segregated). This process may be related to an observer's ability to selectively direct or divide their attention. Second, those signals that are to be integrated should be combined in a way which provides the most reliable joint estimate of their properties. The mathematically optimal way of doing this is to weight each signal's contribution based upon its reliability (Ernst and Banks, 2002; Jacobs, 1999), an operation that is dependent on the observer's ability to accurately estimate the reliability of each piece of information. Bayesian Causal Inference (BCI) provides a normative model for solving these problems in the face of perceptual uncertainty about the world's causal structure, and has been successfully applied to model the behaviour of young, healthy adults (Koerding et al., 2007; Rohe & Noppeney, 2015; Shams & Beierholm, 2010).

Any impact of ageing on multisensory integration might therefore occur at various levels. Deficits in attentional selection and other executive functions may impair older adults' ability to effectively arbitrate between integration and segregation. Increasing age has previously been associated with a reduced ability to separate task-relevant information from distractors, both within (de Fockert, Ramchurn, van Velzen, Bergström, & Bunce, 2009; Helfer, & Freyman, 2008; Panek, Rush, & Slade, 1984; Sekuler, Benett, & Mamelak, 2010) and between (Guerreiro & Van Gerven, 2011; Mevorach, Spaniol, Soden, & Galea, 2016; Poliakoff, Ashworth, Lowe, & Spence, 2006) sensory modalities, as well as with an increased temporal binding window (Diederich, Colonius, & Schomburg, 2008; McGovern et al., 2014). These findings may reflect age-related changes in causal inference, a hypothesis that can be tested both implicitly (by measuring the influence of multiple signals on each other, as in the above paradigms) and explicitly (by directly asking participants whether multiple signals share a source). Within the BCI framework such differences would manifest as a change in the common-source prior, which reflects a participants' estimate of the prior probability of two signals sharing a common source. The reliability-weighted integration of sensory signals may also be impacted by age, due either to actual changes in sensory reliability (Lindenberger & Baltes, 1994; Liu & Yan, 2007; Salthouse, Hancock, Meinz, & Hambrick, 1996) or to impaired reliability estimation. In terms of the BCI model, these would be reflected as differences in sensory variance and relative sensory weighting respectively.

However, the Bayesian causal inference model only makes predictions about final response choices, and therefore implicitly assumes that observers have unlimited time to respond. The real world is not so forgiving: one has limited time to react appropriately to a growling animal. We therefore need to accumulate and evaluate multisensory information over time, and to decide when sufficient evidence has been collected for us to respond effectively. Age-related changes in multisensory integration might therefore also result from how, and for how long, adults of different ages accumulate information before reacting. Indeed, previous research has demonstrated that older adults use a different speed-accuracy trade-off (Starns and Ratcliff, 2010; Salthouse, 1979) and accumulate evidence less efficiently (Madden & Allen, 1995) when responding to unisensory stimuli.

This study combined psychophysics with Bayesian modelling to dissociate the impact of ageing on the various processes and parameters that are critical for multisensory integration. Younger and older adults were characterised and compared in terms of sensory processing, motor speed, selective attention, multisensory integration in an unspeeded ventriloquist paradigm, and multisensory evidence accumulation/speed-accuracy trade-off in a speeded simplified ventriloquist task. Model-free and model-based analyses show that older observers can maintain multisensory spatial localisation performance by sacrificing the speed of their responses.

Methods

Participants

Twenty-four healthy older adults were recruited to the study from a database of participants maintained by the University of Birmingham's School of Psychology. We were unable to contact one following his initial screening session, and a further three were later excluded from analysis: failure to meet inclusion criteria for any one task of the test battery resulted in exclusion from all analyses. Therefore, 20 healthy older adults (six male, mean age = 72.4, *SD* = 5.2, range = 63 – 80 years) were included, as were 20 younger controls (eight male, mean age = 19.1, *SD* = 0.9, range = 17 – 20 years; three exclusions). All reported normal hearing and normal or corrected-to-normal vision. Older participants were also screened using the Montreal Cognitive Assessment; none scored below 25.

Participants gave informed consent prior to the commencement of testing, and were compensated in cash or research credits for their time. The research was given a favourable opinion by the University of Birmingham Ethical Review Committee.

Experimental setup

Participants were seated at a chin rest 130 cm from a sound-transparent projector screen. Behind the screen, at the vertical centre, a shelf held an array of nine studio monitors (Fostex PM04n) spaced horizontally by 7° of visual angle, including a speaker in the middle of the screen. Auditory stimuli were presented via these speakers at approximately 75dB SPL. The locations of the speakers were not known to participants. Images were displayed using a BENQ MP782ST multimedia projector at a total resolution of 1280 x 800. All stimuli were presented using The Psychophysics Toolbox 3 (Kleiner, Brainard, & Pelli, 2007) in MATLAB R2010b running on a Windows 7 PC.

Responses were made using a two-button response pad or optical mouse, and in all cases this was effectively self-speeded; the next trial would not begin until a valid response was made. However, for the cued Eriksen flanker and speeded ventriloquist tasks it was emphasised to participants that they should respond as quickly as possible.

Stimuli

In all but the flanker task, visual stimuli consisted of a 50 ms flash of 15 white (88 cd/m²) dots, each 0.44° of visual angle in diameter, against a dark grey (4 cd/m²) background. Dot locations were sampled from a bivariate Gaussian distribution, with a consistent vertical standard deviation of 5.4°. The horizontal standard deviation of this dot cloud was varied to manipulate the reliability of spatial information, with a wider cloud (expressed in degrees of visual angle) resulting in less reliable stimuli (Rohe & Noppeney, 2015). The auditory stimulus was a simultaneous burst of white noise played from one speaker in the array. Sounds were generated individually for each trial and ramped on/off over 5ms. Participants fixated a central cross (0.22° radius) throughout all tasks.

Experimental design and procedure

Participants completed a battery of tests intended to assess unisensory perception, motor function, selective attention, and audiovisual integration under speeded and non-speeded conditions. All tasks included a short practice session prior to the first full block of trials. The order in which participants took part in these tasks was partially counterbalanced. The first session was used for screening, and thus all participants completed unisensory localisation testing on this occasion. For the other tasks, priority was given to counterbalancing the tasks that were most similar or that were analysed together. Half of participants in each age group completed the continuous response ventriloquist task before the unity judgement, and vice-versa. Similarly, half of participants completed the respond-visual speeded ventriloquist task before the respond-auditory version. Flanker and finger-tapping tasks were completed at various stages and not counterbalanced.

Finger tapping

This task was used to assess participants' motor speed. Participants were instructed to ball their hand into a fist, extending their index finger, and to repeatedly tap a key as quickly as possible for 20 seconds. An on-screen progress bar and countdown provided feedback on performance and time remaining. The task was repeated four times (twice per hand, not including a preceding 10-second practice with each hand).

Auditory left/right localisation

Participants' spatial hearing performance was measured using a forced left/right sound localisation judgement. Individual bursts of white noise were emitted from one of seven locations (-21°, -14°, -7°, 0°, 7°, 14°, or 21°) in a pseudorandom order, and participants told to indicate as accurately as possible via key press whether the sound originated from the left or right half of the screen. This task involved one block of 210 trials (30 per location).

Visual left/right localisation

Participants' ability to localise spatially unreliable visual stimuli was measured using a similar left/right task. Visual stimuli with *SD* = 2° or 25° were pseudorandomly presented centered at one of seven locations (-21°, -14°, -7°, 0°, 7°, 14°, or 21°) and participants again indicated whether each originated from the left or right side of the screen. This task involved two blocks of 210 trials each (total 30 per condition).

Cued Eriksen flanker

A replication of a design conceived by Callejas, Lupiáñez and Tudela (2004), based on and Fan, McCandliss, Sommer, Raz and Posner's (2002) attentional network task, this speeded task assessed three components of attention: alerting, orienting, and executive function. The cueing variable was a 2000 Hz, 50 ms beep, and was present for half of trials. On two thirds of trials, this was followed (after a 400ms delay) by an orienting stimulus: an asterisk, presented for 50 ms either above or below fixation. A standard Eriksen flanker stimulus (Eriksen & Eriksen, 1974), comprising a target arrow flanked by four congruent or incongruent distractor arrows (each of length 1° visual angle, 0.44° separation), was then displayed either above or below fixation and required a speeded left/right direction response. This cued flanker task included 192 trials in a 2 (cueing sound present/absent) x 3 (orienting stimulus congruent/incongruent/absent) x 2 (flankers congruent/incongruent) design. Further details of the paradigm can be found in Callejas et al. (2004) and Fan et al. (2002).

Ventriloquist task – continuous response

This task aimed to evaluate the effect of a synchronous visual stimulus on the perceived location of a sound. It relies on the ventriloquist effect, a perceptual illusion wherein a sound's apparent location can appear to be shifted towards a visual object in another location. In a spatial ventriloquist paradigm, participants were presented with simultaneous auditory and visual stimuli. On each trial these stimuli were independently sampled from five locations (-14°, -7°, 0, 7, or 14°), and could therefore be spatially congruent or incongruent with varying levels of disparity. After a 500ms post-stimulus delay, participants were presented with a mouse-controlled cursor (white, 9° height) whose movement was constrained to the horizontal plane, which they used to report the perceived sound location as accurately as possible. Visual stimuli had three levels of reliability (horizontal *SD* of 2°, 6° or 16°), resulting in a 5 (auditory location) x 5 (visual location) x 3 (visual reliability) design. Participants also responded to a fourth visual reliability level of 0.5°, but this was excluded from analyses due to an error in these stimuli; results were nonetheless similar to those in the 2° condition. Trials were presented psudorandomly in 200-

trial blocks, leading to a total of 600 trials of this task—8 per stimulus combination—as well as separate blocks of unimodal auditory (80 trials) and visual (240 trials) localisation.

Ventriloquist task - common-source judgement

This used the same ventriloquist stimuli as above to provide an explicit measure of participants' common-source judgements. In this case, 500ms post-stimulus the words "same" and "different" appeared respectively above and below the fixation cross. Participants indicated with a button press whether the sound and flash shared a common source. This task was blocked in the same way as the continuous response task, with the same number of trials.

Speeded ventriloquist task

A simplified, speeded version of the above ventriloquist paradigm was used to assess the impact of audiovisual spatial congruence on reaction times and speed/accuracy tradeoff. A visual stimulus with $SD = 5.4°$ was displayed simultaneously with a burst of white noise, centred 14° left or right of a central fixation cross. Stimuli were spatially incongruent on half of trials. Testing was split into two blocks of 160 trials and participants were instructed to indicate with a button press, as quickly as possible, either the location of the sound (respond-auditory) or the dots (respond-visual) while ignoring the other modality. The order of these tasks was counterbalanced between participants. This speeded ventriloquist task therefore employed a 2 (target location) x 2 (distractor location) design, analysed separately for each response modality.

Data analysis

Unimodal localisation

Performance in the unimodal left/right localisation tasks was quantified by fitting psychometric functions to individual participants' responses. In the case of visual stimuli, this was performed separately for each reliability level. The likelihood of a participant responding “stimulus right” was computed for each true stimulus location, and the Palamedes toolbox for MATLAB (Prins & Kingdom, 2009) was used to fit a cumulative Gaussian to these data points. The threshold α and slope β of these functions were allowed to vary freely, while the asymptotes γ and λ were constrained to fall within 0 and 0.05 (Wichmann & Hill, 2001). Measures of left/right discrimination ability (just-noticeable difference at 84% level; JND) and bias (point of subjective equality; PSE) were calculated from the fitted curve (Ernst, 2006). Between-group differences in JND and PSE values were analysed using a two-sample *t*-test in the case of auditory stimuli, and a 2 (reliability) x 2 (age) mixed ANOVA in the case of visual stimuli.

Performance indices for unimodal mouse-cursor localisation tasks were derived by fitting a linear regression (with true stimulus location as the predictor variable) to each participant’s responses. The slope of this line provides an approximate measure for central bias in responses, while the intercept represents biases to the left or right of space. Furthermore, the root mean squared error (RMSE) between the data and the true stimulus locations is a measure of absolute accuracy. Usefully, the above measures of both left/right bias and accuracy produce values in degrees of visual angle. RMSE, L/R bias, and central bias values were entered into two-sample (younger vs. older) *t*-tests in the case of auditory stimuli, and 3 (reliability) x 2 (age) mixed ANOVAs in the case of visual stimuli. Results of the unimodal localisation tasks are summarised in Figure 3.1.

Response time measures

Response time measures were computed as per-condition medians for each participant. It must be noted that response-time data for the ventriloquist sound localisation task should be interpreted with care, as mouse responses are slow and the time required to reach the final response location is heavily dependent on the cursor’s initial position.

The effects of age and congruence were analysed separately for each form of the speeded ventriloquist task (respond-auditory and respond-visual) by entering the above summary statistics into 2 (congruence) x 2 (age) mixed ANOVAs. Similarly, the effects of each of the various attentional factors in the cued flanker task were analysed by entering summary statistics into a 2 (alerting) x 3 (orienting) x 2 (flanker congruence) x 2 (age) mixed ANOVA.

Ventriloquist paradigm

In order to analyse the effect of visual stimuli on reported auditory location we computed crossmodal bias (CMB). This is the distance between the actual and the responded location, scaled by the size of the audiovisual conflict:

$$CMB = \frac{A_{resp} - A_{loc}}{V_{loc} - A_{loc}}$$

A mean of this index was taken for each participant in each condition and entered into a mixed ANOVA (conflict size [pooled over direction] x reliability x age) for model-free analysis. This ANOVA was repeated with median reaction time as the dependent variable.

For the common-source responses, we calculated participants' probability of responding "same-source" for each stimulus combination. As above, these were entered into a mixed ANOVA, and this was repeated for reaction times. We also fitted 3-parameter Gaussian functions (i.e. height, mean, and width) separately for each visual reliability level (Figure 3.3C). The effects of age and visual reliability on each of these parameters were assessed in separate mixed ANOVAs.

Bayesian causal inference model

To better characterise participants' responses, we fitted a Bayesian causal inference (BCI) model of audiovisual perception jointly to the auditory mouse-response and common-source ventriloquist data for each participant (Körding et al., 2007; Rohe & Noppeney, 2015).

The BCI model provides a normative framework for how the brain attempts to estimate the origins of the auditory and visual signals presented simultaneously in this task. The sensory system does not have direct access to their locations and must make a judgement based on noisy sensory inputs. The model therefore accounts for two hypotheses: either the sound and visual stimulus had a common cause, or they were produced by separate sources. In the former case, the location judgement should be same for both stimuli, so the visual estimate can aid the localisation of the auditory (with the relative contribution of each weighted according to their respective sensory reliabilities). In the latter, the stimulus locations should be estimated separately. However, as the sensory system must also contend with uncertainty about whether these signals truly shared a source, some decision strategy must be employed to calculate a final auditory localisation response. We applied each of the three strategies described by Wozny et al., 2010 and colleagues to our data and found that “model averaging” won for both age groups (based on group-level protected exceedance probabilities; Rigoux et al., 2014), so we report parameters fitted with this strategy. When using model averaging, the model gives a response that is a combination of the integrated and segregated estimates, weighted by the estimated probability of them sharing a common source.

Our BCI model had six free parameters: the estimated standard deviation of the auditory (σ_{A}) and visual signals (σ_{V1}, σ_{V2}, and σ_{V3} for each visual reliability level respectively), the prior probability of signals sharing a common source (p_{common}), and the standard deviation of a spatial prior (σ_{P}) that accounts for participants’ tendency to respond towards the centre of space. These parameters were optimised to best fit the data of each participant using a using a nonlinear simplex optimisation function (fminsearch; MATLAB R2016a), the initialisation parameters for which were defined by a grid search. To compare the model fits between age groups, fitted parameters were entered into separate Mann-

Whitney *U* tests, and Bayes factors calculated using the Bayesian Mann-Whitney test as implemented in JASP (JASP Team, 2018; van Doorn, Ly, Marsman, & Wagenmakers, 2018) using the default Cauchy prior. Bayes factors represent the probability ratio of the null (BF_{01}) versus the alternative (BF_{10}) hypothesis, so e.g. $BF_{01} = 3$ indicates that the observed data are three times more likely under the null hypothesis.

Compatibility bias model

We fitted a Bayesian compatibility bias model (Yu, Dayan, & Cohen, 2009) to the speeded respond-auditory ventriloquist data, allowing us to characterise participants in terms of noisy evidence accumulation and response threshold. The idea of this model, as applied in this context, is that participants begin each trial with a prior estimate of the probability that the multisensory stimuli will be spatially congruent. Once the stimuli have been displayed, observers internally accumulate evidence about both the location of the auditory signal and the (in)congruence of the visual distractor. Visual information should be most influential on perceived auditory location at the onset of the trial, when the initial compatibility prior dominates, but this influence decreases as information about the location of each stimulus is accumulated. The process ends when sufficient evidence is accumulated about the location of the auditory stimulus for a decisional threshold to be reached, after which a response is made. The model has five free parameters: the prior probability of stimuli being congruent β; the standard deviation of the auditory and visual evidence accumulation, σ_A and σ_V respectively; the response threshold q; and a non-decision-time parameter t_{nd} that allows for a variable delay between the threshold being reached and a response being given (especially important here, as our finger-tapping task confirmed that the older adults are impaired in raw motor speed). These parameters were adjusted to best fit participants’ data using a Bayesian Adaptive Search optimisation algorithm (BADS; Acerbi & Ma, 2017), with initialisation parameters defined by a grid search. Model fits were again compared between groups by entering each into separate Mann-Whitney *U* tests, and Bayes factors calculated using JASP's Bayesian Mann-Whitney test with default Cauchy prior.

Results

As anticipated, older adults responded significantly more slowly in all tasks. Within-subject task manipulations were generally effective (for example, participants responded more quickly or accurately to congruent stimuli). For consistency across tasks and analyses, we report all effect sizes as eta-squared (η^2), or partial eta-squared (η^2_p) in the case of ANOVAs. Where necessary, ANOVAs were Greenhouse-Geisser corrected for sphericity violations, and *t* tests corrected for unequal variances. Each are indicated by the presence of non-integer *df* values.

Motor speed

The effect of age on motor speed was measured using a speeded finger-tapping task. Data were analysed in terms of the time between finger taps in seconds (inter-tap interval, or ITI) and an overall mean calculated per participant. Younger participants ($M = 0.180$ s, $SD = 0.014$) were significantly faster than their older counterparts ($M = 0.202$ s, $SD = 0.027$), $t(28.772) = -3.269$, $p = .003$, $\eta^2 = 0.220$. No participants were excluded as a result of their performance in this task (all fell within 2 *SD* of their group mean – a conservative threshold).

Unimodal localisation

Results of unimodal localisation tasks are summarised in Figure 3.1. In general, older and younger adults were similarly able to localise auditory stimuli. This was assessed using both left/right and continuous-response approaches. Between-group comparisons of accuracy values indicate that both groups were similarly able to distinguish left from right sounds, $t(38) = 0.278$, $p = .783$, $\eta^2 = 0.002$, and to freely localise auditory stimuli, $t(38) = -1.566$, $p = .126$,

η^2 =.061. Both groups exhibited a rightward bias when localising sounds, though this was larger for older adults in the left/right decision task, $t(32.74) = 2.093$, $p = .044$, $\eta^2 = 0.103$. When providing continuous responses to sound stimuli, older adults exhibited a central bias that was not present in the younger group, $t(38) = 2.44$, $p = .019$, $\eta^2 = 0.136$.

Left/right visual localisation was also similar across groups, with age having no significant main effect or interaction with reliability for either accuracy or bias. Older participants were, however, slightly poorer at freely localising visual stimuli, $F(1, 38) = 5.440$, $p = .025$, $\eta^2_p = .125$, but showed only small biases that were similar to those apparent in the younger group.

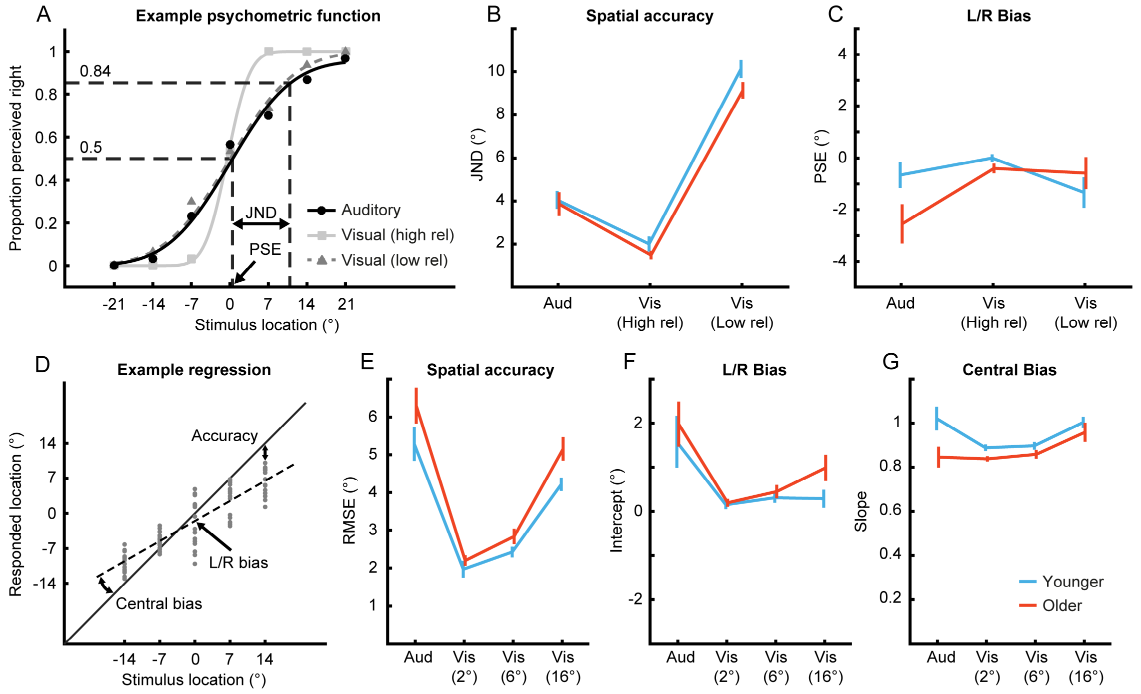

Figure 3.1. Unimodal localisation performance for each age group. (A) Example psychometric function fit to left/right response data for one participant in three conditions, showing calculation of sensitivity (JND) and bias (PSE). (B and C) Mean (± SEM) sensitivity and bias for younger versus older participants. (D) Example continuous-response data for one participant in one condition, with fitted regression line and derivation of accuracy and bias measures. (E, F, and G) Spatial accuracy, left/right bias, and central bias in continuous-response localisation tasks. Data points show group means ± 1 SEM.

Cued Eriksen flanker

The cued flanker task aimed to assess age differences in three attentional networks: alerting, orienting, and executive function. A mixed ANOVA of median response times revealed that participants were significantly faster in congruent flanker trials, $F(1,38) = 383.812$, $p < .001$, $\eta^2_p = .910$, as well as in those with an alerting cue, $F(1,38) = 14.665$, $p < .001$, $\eta^2_p = .278$. Performance was also affected by the presence and spatial congruence of the orienting cue, $F(2,76) = 58.819$, $p < .001$, $\eta^2_p = .608$. However, none of these factors significantly interacted with each other or with the age of participants.

Older adults responded significantly more accurately to this task overall, $F(1,38) = 24.665$, $p < .001$, $\eta^2_p = .394$. Accuracy was lower for trials with incongruent flankers, $F(1,38) = 34.965$, $p < .001$, $\eta^2_p = .479$. This effect was driven primarily by the younger age group: there was a significant interaction between age and flanker congruence, $F(1,38) = 19.948$, $p < .001$, $\eta^2_p = .344$, demonstrating older adults' preference for task accuracy over speed. All other interactions, as well as the main effects of the orienting and alerting cues, had no significant effect on accuracy.

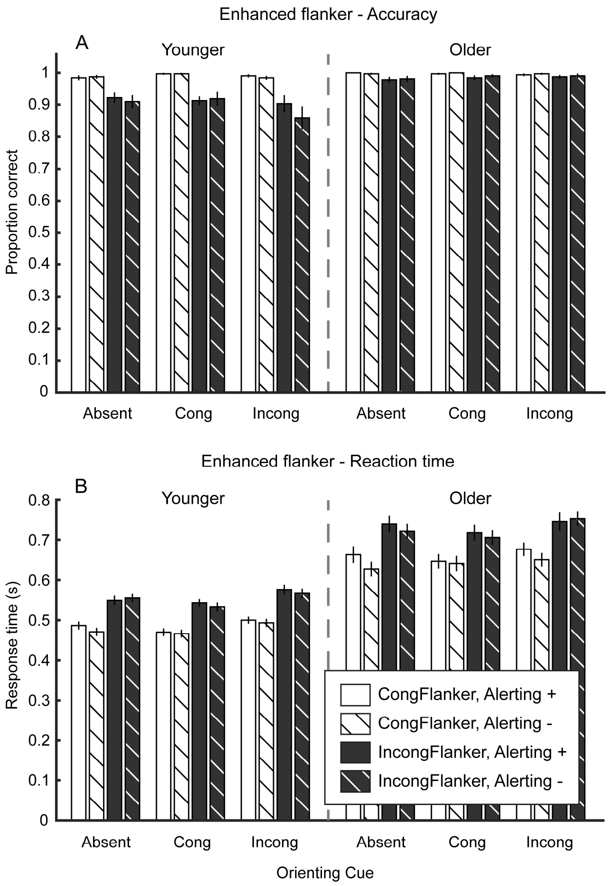

Figure 3.2. Performance on the cued Eriksen flanker task. (A) Flanker task response accuracy in each condition: alerting cue present/absent, orienting cue present/absent, flanker arrows congruent/incongruent. (B) Group means of participants' median response times. Error bars show ± 1 *SEM*.

Ventriloquist paradigm

Auditory localisation responses

In order to quantify the effect of a simultaneous visual stimulus on the perceived location of a sound (Figure 3.3A), we analysed reported sound location in terms of the relative effect of the visual signal ($(A_{resp} - A_{loc}) / (V_{resp}) - (V_{loc})$). As anticipated, this effect was stronger at smaller audiovisual conflict sizes, $F(2.100, 79.800) = 108.891$, $p < .001$, $\eta^2_p = .741$, and when the visual stimulus was more reliable, $F(2, 76) = 46.048$, $p < .001$, $\eta^2_p = .548$. Conflict size also interacted with stimulus reliability, $F(4.181, 158.891) = 27.370$, p

$< .001$, $\eta^2_p = .419$. Crucially, however, age had no overall effect on participants' responses, nor did it interact with conflict size or stimulus reliability. In short, older participants did not differ from the younger group in their susceptibility to the ventriloquist effect.

Aside from a main effect of age, $F(1, 38) = 8.047$, $p = .007$, $\eta^2_p = .175$, response times did not differ between conditions (Figure 3.3B). This is unsurprising, as mouse movements are far more variable (and take much longer) than button presses, so any small effects of condition are likely to be lost.

Common-source judgements

Common-source judgements provide an explicit measure of participants' tendency to integrate or segregate audiovisual stimuli under various conditions. Responses are analysed in terms of the probability of responding "same source" in each condition (Figure 3.3C). As expected, the probability of integrating was higher when the auditory and visual stimuli were closer together, $F(1.731, 65.788) = 352.795$, $p < .001$, $\eta^2_p = .903$, and when visual stimuli were more reliable, $F(1.396, 53.061) = 13.159$, $p < .001$, $\eta^2_p = .257$. These factors also interacted, $F(4.950, 188.102) = 10.429$, $p < .001$, $\eta^2_p = .215$, suggesting audiovisual disparity was a more informative cue to integration when stimulus reliability was higher. Age, however, had no effect on these probabilities.

Fitting three-parameter Gaussians to these data separately for each reliability level further confirms that visual reliability modulates integration, indicated by a significant effect of reliability on the height parameter, $F(1.423, 54.086) = 41.871$, $p < .001$, $\eta^2_p = .524$. Age again had no effect on any of the fitted parameters in any condition. This demonstrates that older and younger adults integrate audiovisual stimuli under similar conditions and, combined with the above localisation data, further confirms that the two age groups ultimately respond similarly to complex multisensory stimuli given sufficient time.

However, reaction times in the common-source judgement task reveal a number of age-related effects, despite participants receiving no instruction to respond quickly. Across both age groups, participants responded faster for the common-source judgements that were easier to perform: response times were faster for stimuli that were perfectly congruent or extremely incongruent than for those with small audiovisual disparities, $F(1.819, 69.108) = 11.617$, $p < .001$, $\eta^2_p = 234$, and participants responded more slowly when the visual stimulus was less reliable, $F(1.560, 59.274) = 10.926$, $p < .001$, $\eta^2_p = .224$. An interaction between these factors was also apparent, $F(5.878, 223.345) = 4.383$, $p < .001$, $\eta^2_p = .103$. Age interacted significantly with both audiovisual disparity, $F(1.819, 69.108) = 3.633$, $p = .036$, $\eta^2_p = .087$, and stimulus reliability, $F(1.560, 59.274) = 5.697$, $p = .010$, $\eta^2_p = .130$. The three-way disparity x reliability x age interaction was also significant, $F(5.878, 223.345) = 4.108$, $p = .001$, $\eta^2_p = .098$. This suggests that older adults took more time to process complex incongruent stimuli, but ultimately responded in a similar way to the younger group.

BCI model

The Bayesian causal inference model was fitted jointly to each participant's auditory localisation and common-source judgement responses. Table 3.1 summarises the parameters of these fits, including nonparametric significance tests of group differences and corresponding Bayes factors. Significant differences in the fits between groups would be indicative of age-related changes in multisensory processing: a higher value for P_{common} would, for example, suggest reduced or impaired segregation. Differences in σ_A and $\sigma_{V1\text{-}V3}$ would imply changed estimation of sensory reliability. In agreement with the model-free analyses above, however, there was no evidence of a between-group difference for any of the fitted parameters. The fitted parameters vary considerably between participants, but the size of this variability is similar across the age groups, and the averaged group fits are qualitatively good and very similar between groups (Figure 3.3A and 3.3C, dashed lines). These results are summarised in Table 3.1.

Table 3.1.

Summary statistics and separate Mann-Whitney U tests of fitted BCI parameters.

		P_{common}	σ_{prior}	σ_{auditory}	σ_{visual1}	σ_{visual2}	σ_{visual3}
Younger	*M*	0.389	41.455	8.464	2.247	4.659	15.408
	SD	0.128	29.117	3.740	5.529	5.907	10.588
Older	*M*	0.403	35.711	9.174	0.996	4.300	17.780
	SD	0.118	25.991	4.241	2.894	7.866	16.662
Mann-Whitney	*U*	188	172	172	158	154	200
	p	.745	.449	.449	.256	.213	> .999
	η^2	.003	.014	.014	.032	.039	0
Bayes factors	BF_{10}	0.342	0.441	0.378	0.530	0.475	0.321
	BF_{01}	2.924	2.694	2.266	1.887	2.106	3.120

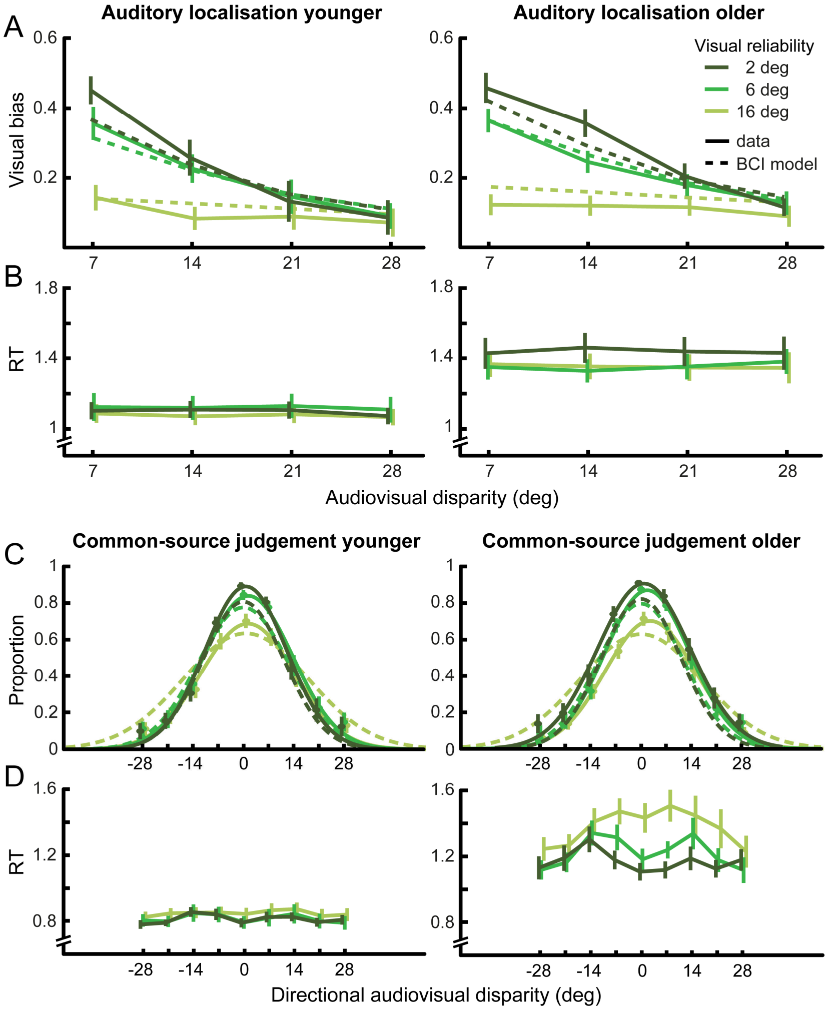

Figure 3.3. Behavioural performance and BCI model predictions for ventriloquist tasks, separated by age group. (A) Magnitude of visual bias ($[A_{resp} - A_{loc}] / [V_{loc} - A_{loc}]$) in auditory localisation task, shown as a function of audiovisual disparity (*x*-axis, pooled over direction) and visual reliability (colour coded). (B) Reaction times in auditory localisation task. (C) Proportion “same-source” responses in common-source judgement task, as a function of audiovisual disparity and visual reliability. Dots show actual values; curves are fitted Gaussians. (D) Reaction times (pooled over response) in common-source judgement task. All behavioural data points show group means ± 1 *SEM*.

Speeded ventriloquist paradigm

Respond auditory

The speeded L/R ventriloquist paradigm assessed the effects of multisensory stimuli on participants’ responses under speeded conditions, providing more suitable data for reaction time analysis. When responding to the location of the auditory stimuli, participants reacted overall more quickly when the visual distractor was spatially congruent with the auditory target, $F(1,38) = 63.314$, $p < .001$, $\eta^2_p = .625$. Importantly, we also found an interaction between the effects of age and congruence, $F(1,38) = 12.635$, $p = .001$, $\eta^2_p = .250$, indicating that older participants' response times were more affected by the spatial (in)congruence of the stimuli.

Participants responded less accurately in the respond-auditory task when the visual distractor was spatially incongruent, $F(1,38) = 46.970$, $p < .001$, $\eta^2_p = .553$, but this was not significantly affected by age and the factors did not interact.

Respond visual

Participants were again faster at responding to targets with a spatially congruent distractor in this task, $F(1,38) = 33.705$, $p < .001$, $\eta^2_p = .470$, but age did not interact with this effect.

In comparison to the respond-auditory task, response accuracy was less affected by stimulus congruence, $F(1,38) = 4.394$, $p < .043$, $\eta^2_p = .104$. In this case, however, response accuracy was impacted by age, with older participants proving slightly more accurate than their younger counterparts in both conditions, $F(1,38) = 5.163$, $p = .029$, $\eta^2_p = .120$. The absence of an interaction between age and congruence, $F(1,38) = 1.099$, $p = .301$, $\eta^2_p = .028$, however, suggests that this effect was independent of the task manipulation. See Figure 3.4B and C for accuracy and RT plots.

Compatibility bias model

A compatibility bias model was fitted to participants' respond-auditory task responses to characterise them in terms of evidence accumulation and decision thresholds. Fitted parameters were compared using separate Mann-Whitney *U* tests and a Bayesian equivalent;

see Table 3.2 for a summary of results. Confirming the findings of the BCI model, the age groups did not differ in their prior tendency to integrate multisensory stimuli, characterised in this case by the parameter β. The groups also did not differ in the reliability of visual evidence accumulation σ_{visual}. The remaining three parameters were significantly different between the groups. Non-decision time t_{nd}, which captures the time between a decision is made and the response given (and is therefore closely linked to motor speed), was significantly higher for the older age group. Older adults also set their decision threshold q significantly higher, requiring more evidence before deciding on a response. Finally, and crucially, older adults were significantly noisier in accumulating auditory evidence, reflected as higher σ_{auditory} values. This indicates that it takes older participants longer than their younger counterparts to reach any given level of certainty about the location of an auditory stimulus. See Figure 3.4A for an illustration of the model.

Table 3.2.

Summary statistics and separate Mann-Whitney U tests of fitted compatibility bias parameters.

		σ_{auditory}	σ_{visual}	β	q	t_{nd}
Younger	*M*	0.887	0.681	0.804	0.918	0.234
	SD	0.397	0.999	0.103	0.046	0.058
Older	*M*	1.424	0.500	0.772	0.943	0.358
	SD	0.857	0.387	0.140	0.072	0.068
Mann-Whitney *U*	*W*	291.0	218.0	195.5	291.0	355.0
	p	**0.013**	0.640	0.914	**0.013**	**< .001**
	η^2	0.151	0.006	0	0.151	0.439
Bayes factors	BF_{10}	**6.172**	0.313	0.329	**2.531**	**162.239**
	BF_{01}	0.162	3.196	3.038	0.395	0.006

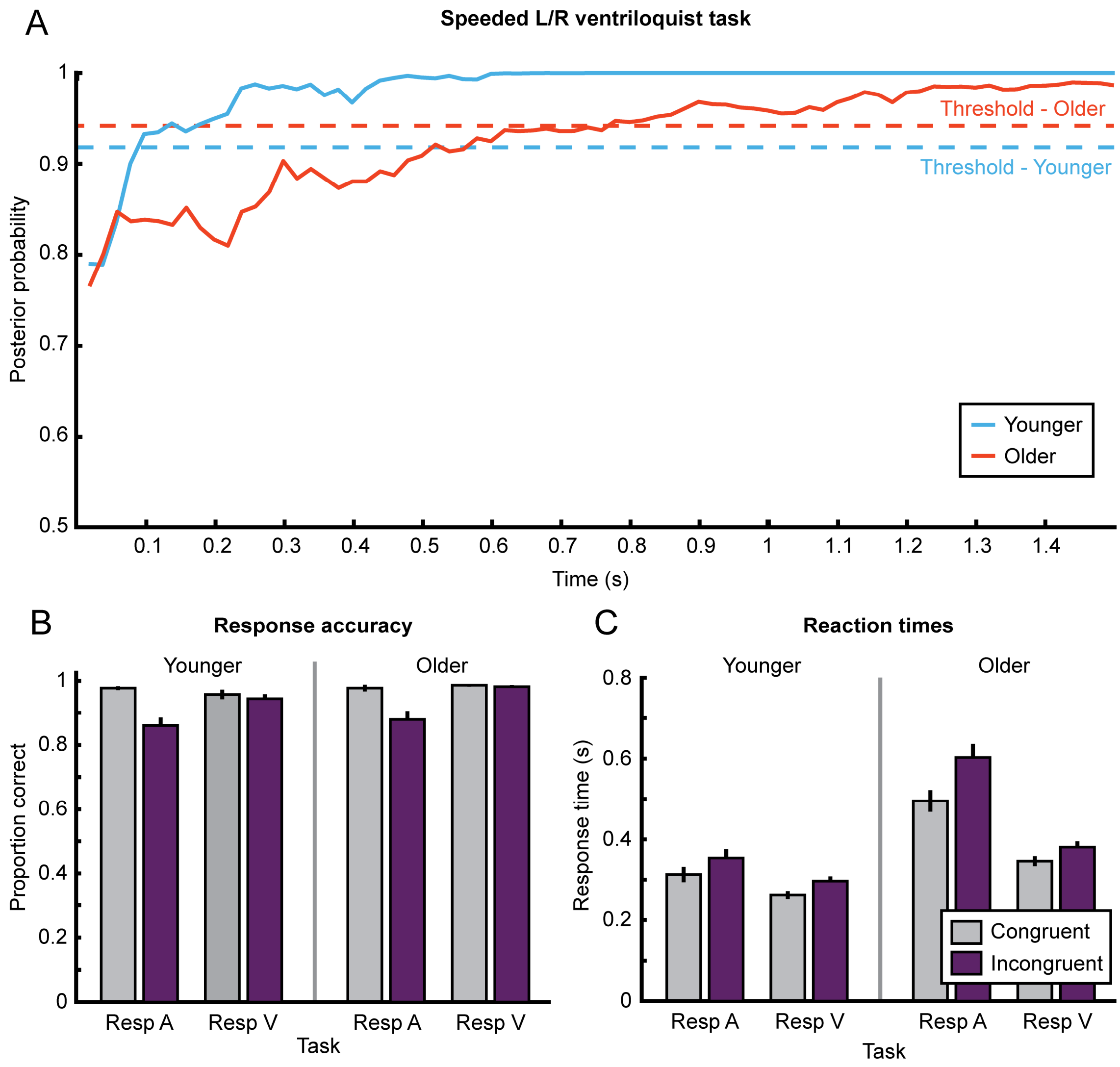

Figure 3.4. Results of speeded left/right ventriloquist task. (A) Illustration of group differences in evidence accumulation in one experimental condition (auditory left, visual right) of the respond-auditory task, according to compatibility bias model fits. Solid lines show the posterior probability that the auditory stimulus is on the left (summed over independent [two sources] and joint [one source] estimates) plotted over time. Dashed lines show decision thresholds. Older adults can be seen to accumulate evidence more noisily, and to have a higher decision threshold. This plot shows group means of simulations run once per participant using their best-fit parameters, to highlight group differences while also illustrating the randomness in the accumulation. (B and C) Response accuracy and reaction times in respond-auditory and respond-visual tasks, separated by spatial congruence (i.e. pooled over left and right). Error bars show ± 1 *SEM*.

Discussion

This study aimed to quantify the effects of ageing on the processes underlying multisensory integration. Using model-free analyses, we demonstrated that younger and older adults performed comparably across a number of domains: unisensory localisation (using left/right decisions and continuous localisation), attentional processing (measured via the cued Flanker task), and multisensory integration (including both explicit and implicit causal inference) were all similar across age groups. Likewise, model-based analysis using a Bayesian causal inference model fitted to response choices revealed no significant group differences in sensory variances or spatial priors. The common-source prior, which quantifies observers' tendency to bind multisensory stimuli, was similarly unaffected by age.

Despite this multitude of similarities, older participants were found to be substantially slower in all tasks. A large component of this difference must be age-related decline in motor speed, as highlighted by the results of our finger-tapping measure. However, in terms of reaction times, age also interacted with task manipulations on both speeded and non-speeded multisensory tasks, suggesting that older adults may sacrifice speed to preserve performance under challenging conditions. By considering response threshold and non-decision time (i.e. motor delay) as separate parameters, as well as quantifying noise in evidence accumulation, the compatibility bias model (fitted to responses in a speeded ventriloquist task) allowed us to probe the computational mechanisms underlying age differences in response speed. Between-group comparisons of fitted parameters confirmed that older adults did exhibit longer non-decision time, but also showed that they set their response threshold slightly higher and, crucially, were noisier in accumulating sensory evidence than the younger age group. The fits also confirmed the finding from the BCI model that the age groups did not differ in their prior tendency to bind signals (P_{common} for BCI, β for compatibility bias). Taken together, these results suggest that older adults successfully compensate for extra sensory noise by accumulating internal evidence for longer.

Our results align well with previous research. Response slowing is a well-documented, robust effect of increasing age (Der & Deary, 2006; Fozard, Vercruyssen, Reynolds, Hancock, & Quilter, 1994; Salthouse, 2000), as is the tendency to prioritise accuracy over speed on reaction time tasks (Hertzog, Vernon, & Rypma, 1993; Rabbitt, 1979; Salthouse, 1979; Smith & Brewer, 1995). Prior studies have attempted to explicitly model the effects of age on evidence accumulation and response thresholds using a variety of models and (unisensory) tasks. Madden & Allen (1995) found evidence for age-related differences in evidence accumulation by simply assessing the relationship between reaction time and accuracy in response to a visual classification task. Ratcliff (2008) later assessed older and younger adults' performance on a numerosity judgement task, using standard speed or accuracy instructions as well as a response signal approach in which participants may only respond after a cue presented at a set interval. Diffusion models fitted separately to each response type agreed that age had significant effects on both decision threshold and non-decision time. Starns and Ratcliff (2010) reached similar conclusions using data from brightness judgement, letter discrimination, and recognition memory tasks, and noted that older adults were less able to adjust their threshold in response to varying task instructions. Finally, Forstmann et al. (2011) confirmed this finding, as well as demonstrating age differences in the accumulation process, by fitting a linear ballistic accumulation model to the results of a motion discrimination task.

Our findings also agree with, and can help to account for, previously shown age differences in reaction times in the presence of cross-sensory distractors. Guerreiro and Van Gerven (2011), for example, found that older adults reacted more slowly to a visual *n*-back task when an auditory distractor was present, while younger adults were unaffected. Similarly, Mevorach et al. (2016) showed that visual distractors had a far greater impact on older adults' response times than they did for the younger group. Noisier evidence accumulation and higher response thresholds may help to explain these effects.

Our finding that older adults are no more influenced by the ventriloquist effect than younger adults (in terms of final responses) contrasts with the those of other multisensory illusions: responses to the double-flash (DeLoss et al., 2013; McGovern et al., 2014; Setti et al., 2011), McGurk-MacDonald (Sekiyama et al., 2014; Setti et al., 2013), and stream/bounce (Roudaia et al., 2013) illusions are all substantially influenced by age. One explanation for this difference is the presence of a strong temporal component in each of these illusions that is absent in the ventriloquist paradigm. While age has been shown to have limited impact on both unisensory and bisensory spatial localisation (Dobreva, O'Neill and Paige, 2012), older adults are known to be impaired in a variety of timing perception tasks including temporal order judgement (Ulbrich, Churan, Fink, & Wittmann, 2009) and gap detection (Snell, 1996), possibly due to a decrease in neural representational precision (Frisina & Walton, 2006; Walton, Frisina, & O'Neill, 1998). Temporal fine structure perception, a skill crucial to speech perception in adverse environments (Füllgrabe, Moore, & Stone, 2015), has also been shown to deteriorate with age (Hopkins & Moore, 2011).

We must, however, acknowledge that the intense, multi-session nature of this research may have led to some bias in our selection of older adults: participants who voluntarily sign up to a research database and agree to a multiple-day study will probably perform above average for their age, and it is possible that a different pattern of results may have emerged had we recruited from (for example) a retirement home. Despite this, our older participants did show clear age-related slowing of responses, and the dissociation between reaction time measures and final responses in our results provides strong support for the conclusion that older adults attempt to preserve response accuracy in the presence of noisier multisensory stimulus representations by internally accumulating evidence and evaluating the stimuli for longer.

CHAPTER 4: CORTICAL HIERARCHIES FOR MULTISENSORY INTEGRATION ARE PRESERVED IN HEALTHY OLDER ADULTS

Samuel A. Jones, Susan Francis, Stephen Mayhew, and Uta Noppeney

The efficient integration of multisensory stimuli is central to our ability to navigate in the world. To effectively interpret the rich array of signals that are continually received by our sensory systems, we must identify those which have a common source and integrate them while processing signals from other sources separately. To extract the most information from signals that do share a source, each one's influence in the final estimate should be weighted by its relative reliability. Young adults have been shown to perform statistically optimally at this task in a wide range of situations (Alais & Burr, 2004; Ernst & Banks, 2002; Jacobs, 1999; Koerding et al., 2007; Shams & Beierholm, 2010), and multivariate analysis of neural responses to spatially incongruent audiovisual signals has revealed that the process takes place across a cortical hierarchy that encompasses primary sensory areas as well as higher-level regions such as intraparietal sulcus and planum temporale (Rohe & Noppeney, 2015, 2016, 2018).

However, normal healthy ageing leads to a variety of sensory and cognitive changes, including loss of sensory acuity (Faubert, 2002; Frisina & Walton, 2006; Liu & Yan, 2007), impaired attentional and working memory processes (Gazzaley, Cooney, Rissman, & D'Esposito, 2005; Hasher & Zacks, 1988; Zanto & Gazzaley, 2014), and reduced processing speed (Eckert, Keren, Roberts, Calhoun, & Harris, 2010; Salthouse, 1991, 1996, 2000). There is some evidence that these factors affect how older people respond to multisensory stimuli. Susceptibility to the sound-induced flash illusion has consistently been shown to increase with age (DeLoss, Pierce, & Andersen, 2013; McGovern, Roudaia, Stapleton, McGinnity, & Newell, 2014; Setti, Burke, Kenny, & Newell, 2011), and the McGurk effect is also stronger in older adults (Sekiyama, Soshi, & Sakamoto, 2014; Setti, Burke, Kenny, & Newell, 2013). However, older adults also seem to receive the same or, possibly even more, reaction time benefit from congruent multisensory stimuli as younger people (Diaconescu, Hasher, &

McIntosh, 2013; Laurienti, Burdette, Maldjian, & Wallace, 2006; Mahoney, Li, Oh-Park, Verghese, & Holtzer, 2011), suggesting that redundant information may help to compensate for the impairments described above.

We recently conducted a behavioural study that compared the optimality of older and younger adults' responses to conflicting audiovisual spatial cues (Jones, Beierholm, and Noppeney, 2018; Chapter 3 of this thesis). This study revealed both similarities and differences between the age groups. The groups performed comparably on unimodal stimulus localisation tasks in both auditory and visual modalities, and older adults' final localisation and common-source judgement responses closely matched those of younger participants and were well-predicted by a statistically optimal Bayesian causal inference model. Despite this, older participants' reaction times were significantly slower across a range of tasks and interacted with difficulty: older adults took longer to respond to more challenging stimuli. A compatibility bias model fitted to participants' responses to a speeded ventriloquist-type task revealed that older adults needed to spend more time accumulating evidence from the multisensory signals before reaching the same degree of certainty as the younger group.

One possible interpretation of this finding is that older adults rely on extra compensatory processes because their representation of, and/or ability to efficiently process, multisensory stimuli is degraded. Age differences in neural activity provide some evidence for such compensatory activity in response to a variety of tasks (Cabeza, 2002; Davis, Dennis, Daselaar, Fleck, & Cabeza, 2008; Park and Reuter-Lorenz, 2009; Grady, 2012). Imaging research specifically addressing the interaction between and age and multisensory integration is limited, however. In some cases, electrophysiological work has demonstrated decreased responses to multisensory stimuli in older adults (Setti et al., 2011; Stephen, Knoefel, Adair, Hart, & Aine, 2010), while others have found increased neural activity that has been interpreted as compensatory (Diaconescu, Hasher, & McIntosh, 2013; Winneke & Phillips, 2011; Zou, Chau, Ting, & Chan, 2017). Functional MRI work by Townsend, Adamo, and Haist (2006) showed that, when performing an audiovisual attention-switching task, older participants consistently recruited a wide network of frontoparietal regions that was only activated by younger groups in the most demanding conditions.

However, recent work by Morcom and colleagues (Morcom, Henson, & Cam-CAN, 2018; Morcom & Jonson, 2015) has emphasised the need for caution in the interpretation of activation changes between age groups. An increase in cortical activity is not necessarily compensatory and may instead reflect inefficiency or increased effort. More sensitive approaches, such as multivariate decoding methods that specifically address the type and quality of information encoded in a region, are therefore needed to establish the nature of any age differences in task-related neural activity.

The present fMRI study used a modified ventriloquist paradigm to address the question of cortical mechanisms underlying multisensory integration in the ageing brain. Healthy younger and older participants were presented with audiovisual signals at various degrees of spatial conflict and required to selectively attend to (and report) the location of the auditory component of the stimulus. We applied standard mass-univariate analysis alongside two types of multivariate decoding to establish whether older adults relied on compensatory processes to preserve their behavioural responses to these stimuli, and to test whether representations of stimuli across the cortical hierarchy differed between age groups.

Methods

Participants

Twenty younger and 29 older adults were initially recruited from participant databases for a behavioural screening session. Two older adults were excluded from the study due to the presence of MRI contraindications, three failed to score above 24 on the Montreal Cognitive Assessment (Nasreddine et al., 2005), and one reported taking antidepressant medication. A further seven older, and three younger, adults were excluded for insufficient gaze fixation in the behavioural task (see below for details). One younger participant could not be contacted following the behavioural session. Therefore, 16 younger (mean age = 24.19, *SD* = 4.56, range = 19 – 32 years, 10 female) and 16 healthy older (mean age = 70.75, *SD* = 4.71, range = 65 – 82 years; 12 female) adults took part in all three experimental sessions. All included participants had normal or corrected-to-normal vision, reported no hearing impairment, and were able to distinguish left from right sounds with a just-noticeable difference of below 7°. All participants gave informed consent to take part in the study, which was given a favourable ethical opinion by the University of Birmingham Ethical Review Committee.

Stimuli

Visual stimuli consisted of a cloud of 20 white dots (diameter of 0.4° visual angle), whose locations were sampled from a bivariate Gaussian distribution with a standard deviation of 2.5° in both directions, presented on a black background. The auditory stimulus was a burst of white noise, convolved with generic (i.e. non-individualised) head-related transfer functions (HRTFs) to simulate spatial locations. These HRTFs (MIT Media Lab) provide both monaural and binaural spatial filtering signals as cues to the location of a sound's source. Sounds were presented with a 5ms on/off ramp.

Design and procedure

In a spatial ventriloquist paradigm, participants were presented with synchronous auditory and visual signals and asked to report the location of the sound. On each trial, the auditory signal could originate from any one of four possible spatial locations (-15°, -5°, 5°, or 15° azimuth). For any given auditory location, a synchronous visual signal could be: presented centred at the same spatial location (congruent trial), presented at the symmetrically opposite location (incongruent trial), or absent (unimodal trial). This resulted in a 4 (auditory location) x 3 (congruence) factorial design (see Figure 4.1A for a visual outline).

On each trial, the auditory signals were presented synchronously with the visual (if present) for 80ms with an SOA of 2.3 s. Participants were instructed to indicate the perceived location of each sound by pressing one of four spatially corresponding buttons with their right hand, and to fixate a central cross (0.75° diameter) throughout the experiment.

Each five-minute experimental run consisted of 120 trials, with every condition repeated exactly 10 times. To increase the efficiency of the design, conditions were arranged into mini-blocks of one, two, three, and four trials within each run, the order of which was randomised. These mini-blocks were interspersed with 6.9 s fixation periods approximately every 20 trials. Eleven runs were acquired across two separate scanning sessions for a total of 110 trials per condition.

Experimental Setup

Stimuli were presented using Version 3 of the Psychophysics Toolbox (Kleiner, Brainard, & Pelli, 2007), running on MATLAB 2014b on an Apple Macbook. Auditory stimuli were presented at approximately 75 dB SPL through Optime 1 electrodynamic headphones (MR Confon). Visual stimuli were back-projected by a JVC DLA-SX21E projector onto an acrylic screen, viewed via a mirror attached to the MRI head coil. The total viewing distance from eye to screen was 68cm. Participants responded using infrared response pads (Nata Technologies) held in the right hand.

Behavioural testing session

Prior to MRI acquisition, participants underwent one lab-based session of behavioural training and screening. First, they completed a simple left/right forced-choice localisation task to ensure they were able to accurately localise HRTF-transformed auditory stimuli. On each trial an auditory stimulus was presented randomly from one of ten locations between -15° and 15° azimuth. The task was to indicate whether they perceived the sound as coming from the left or right by pressing one of two buttons.

Participants then completed a localisation training session to learn the mapping between the four buttons and the corresponding auditory locations (-15°, -5°, 5°, and 15°) used in the main ventriloquist task. On each trial, a sound was presented randomly from one of the four locations and participants attempted to locate it with a button press. Feedback was provided after each response: correct responses were rewarded with a green square; incorrect responses resulted in a red square, followed by a green square in the correct location. Participants completed up to five 20-trial blocks, stopping early if localisation accuracy reached 90% in any block.

Finally, participants completed two blocks of the ventriloquist task used during the scanning session. During these blocks a recording of an fMRI sequence was played over speakers at a level that approximately matched that experienced in the scanner (after adjustment for headphone attenuation). To minimise the possibility of eye movement confounds in the fMRI data, we also screened participants for their ability to maintain central fixation during this task. Throughout the two blocks, participants' eye movements were recorded via a Tobii EyeX eye tracker. A custom MATLAB script was used to remove blinks and identify saccades (defined as four or more consecutive samples greater than two degrees from the fixation cross). For each participant, the peak location (i.e. furthest from fixation) of every recorded saccade was entered as data into a linear regression analysis, with per-trial visual stimulus locations as the predictor variable. Any participant for whom the stimulus location significantly predicted peak saccade location was not invited back for the MRI sessions. In this way, participants with stimulus-driven saccades were excluded from the study (seven older adults; three younger).

Behavioural analysis

A psychometric function (cumulative Gaussian) was fitted to the left/right sound localisation responses using the Palamedes toolbox for MATLAB (Prins & Kingdom, 2009), and quantified in terms of the point of subjective equality (PSE; a measure of bias) and just-noticeable difference (JND; a measure of accuracy) at the 84% level (Ernst, 2006). These were entered into separate independent-samples *t*-tests to compare performance between groups, and a JND of greater than 10° was specified as an exclusion criterion (all participants met this requirement).

To characterise participants' responses to the ventriloquist stimuli we calculated their mean localisation choice for each location under each condition (unimodal, congruent, and incongruent). Audiovisual response data were entered into a 3 (unimodal / congruent / incongruent) x 4 (auditory location) x 2 (young / old) mixed ANOVA. A group-level correlation of mean responses against true stimulus locations was also conducted to verify that participants were able to accurately localise the sounds.

MRI data acquisition

A 3T Philips Achieva MRI scanner was used to acquire T_1-weighted anatomical images and T_2*-weighted BOLD EPI (gradient echo, SENSE factor of 2, TR = 2800 ms, TE =

40 ms, flip angle = 90°, FOV = 192×192x114 mm², 38 slices, voxel size = 2.5×2.5×2.5 mm^3 with a 0.5 mm interslice gap).

Scanning of each participant took place over two separate one hour-sessions. Over these sessions, eleven task runs of 115 volumes each was acquired. Each session also involved a further 115-volume resting-state run, during which participants were instructed to fixate a central cross. Four dummy volumes were acquired and discarded before each run to allow T_1 magnetisation to reach equilibrium.

fMRI data analysis

We took three approaches to analysis of the fMRI data. First, we investigated task and age-related effects using a conventional second-level mass-univariate general linear model (GLM). Based on functionally-defined regions from this analysis, we then conducted multivariate Bayesian decoding (MVB) analyses to determine the amount of stimulus-relevant information held in the areas activated more by older adults in comparison to the areas activated by both age groups. Finally, we used a multivariate support vector regression (SVR) analysis to quantify differences in the type and quality of information stored in a series of a-priori defined regions within the audiovisual cortical hierarchy. Analyses were conducted on all trials, irrespective of the behavioural response, though these were included as regressors of no interest in some cases (see below for details).

MRI data were analysed in SPM12 (Friston et al., 1994). Each participant’s functional scans were slice-timing corrected, realigned/unwarped to correct for movement, and coregistered to the anatomical scan. For the analyses conducted in native space (SVR and MVB), these data were then spatially smoothed with a Gaussian kernel of 3mm FWHM. For group-level GLM analyses, the slice-timing-corrected and realigned images were first normalised into Montreal Neurological Institute (MNI) space by applying a deformation field produced by running SPM's unified segmentation-normalisation on the anatomical image. The normalised functional images were then resampled to a voxel size of $2mm^3$ and spatially smoothed with a Gaussian kernel of 8mm FWHM. The following steps were conducted separately on both native-space and MNI-transformed data. Each voxel's time series was high-pass filtered with a cutoff of 1/128Hz. Stimuli were modelled as events of zero duration, and their onsets convolved with a canonical haemodynamic response function and its temporal derivative. As well as regressors for each of the conditions in the factorial design, motion parameters were included in the model as nuisance regressors to account for residual effects of movement. The GLM also included the button response locations (and temporal derivative) as a parametric regressor of no interest to limit the influence of motor responses in the group-level GLM and MVB analyses. This regressor was not included for the support vector regression analysis as doing so made the predictions noisier but revealed no new age differences.

To account for age-related changes in vascular reactivity that may confound any between-group BOLD activation differences, we applied resting state fluctuation amplitude (RSFA) correction (Kannurpatti & Biswal, 2008; Tsvetanov et al., 2015) to our parameter estimates in the MNI-normalised data. The basic idea of this approach is to use signal variability in each voxel as a measure of vascular reactivity, and to correct BOLD activations based on this measure. Resting-state data were first preprocessed in the same way as the task data. We then applied the following additional steps: wavelet despiking (Patel et al., 2014) to remove motion artefacts; linear and quadratic detrending; regression of mean white matter and cerebrospinal fluid signal; regression of movement parameters (and first derivatives); and bandpass-filtering at 0.01-0.08Hz to maximise the contribution of physiological factors to the signal fluctuation. The standard deviation of the remaining signal at each voxel was calculated to create the final RSFA map, produced separately for each scanning session. We then divided every subject-level parameter estimate by the relevant RSFA value.

Group-level GLM

We implemented a random-effects analysis at the group level by entering participants' parameter estimates for each stimulus condition (averaged across the 11 runs) as data in a new GLM, then testing specific effects with *t* contrasts. Unless otherwise stated, we report results at $p < .05$ with whole-brain FWE-correction at the voxel level. First, we verified whether the design was generally effective across age groups by testing for the main effect of stimulus incongruence ($\text{AVIncongruent}_{\text{All}} > \text{AVCongruent}_{\text{All}}$). We then examined age differences in general task-related activation, pooled across conditions ($[\text{All}_{\text{Older}} > \text{Baseline}_{\text{Older}}] > [\text{All}_{\text{Younger}} > \text{Baseline}_{\text{Younger}}]$, and $([\text{All}_{\text{Younger}} > \text{Baseline}_{\text{Younger}}] > [\text{All}_{\text{Older}} > \text{Baseline}_{\text{Older}}])$. Finally, we searched for interactions between age and stimulus congruence ($[(\text{AVIncong}_{\text{Older}}) > (\text{AVCong}_{\text{Older}})] > [(\text{AVIncong}_{\text{Younger}}) > (\text{AVCong}_{\text{Younger}})]$, and $[(\text{AVIncong}_{\text{Younger}}) > (\text{AVCong}_{\text{Younger}})] > [(\text{AVIncong}_{\text{Older}}) > (\text{AVCong}_{\text{Older}})]$).

Multivariate Bayes

For each participant, several multivariate Bayes (MVB) models were fitted to assess whether areas with increased cortical activation in older adults also held more information about the stimuli, i.e. whether this activity was compensatory. MVB decoding, as implemented in SPM12, uses the same design matrix as the univariate GLM but reverses the mapping. In a standard GLM analysis, the details of an experimental design attempt to predict the physiological (BOLD) outcome. Conversely, MVB uses a hierarchical parametric empirical Bayes model to predict a target variable of interest using many physiological variables (BOLD activity across numerous voxels). The target variable, for example congruent versus incongruent stimuli, is specified as a contrast in the GLM; all other regressors are modelled out as confounds. MVB attempts to find the pattern of voxels that best predicts the target variable while also optimising the set size in terms of free energy. For further details on the implementation of MVB see Friston et al. (2008). We specified a sparse spatial prior, as this has been previously demonstrated to best describe the distribution of activations across a variety of tasks (Chadwick, Bonnici, & Maguire, 2014; FitzGerald, Friston, & Dolan, 2012; Hulme, Skov, Chadwick, Siebner, & Ramsøy, 2014; Morcom & Friston, 2012; Morcom, Henson, & Cam-CAN, 2018).

The three ROIs from which voxel patterns were extracted for the MVB models were specified functionally based on group-level GLM contrasts: $(\text{All}_{\text{Older}} > \text{Baseline}_{\text{Older}}) > (\text{All}_{\text{Younger}} > \text{Baseline}_{\text{Younger}})$, henceforth referred to as O>R, representing the regions that older adults were recruiting more than younger adults while performing the task; the conjunction $(\text{All}_{\text{Younger}} > \text{Baseline}_{\text{Younger}}) \cap (\text{All}_{\text{Older}} > \text{Baseline}_{\text{Older}})$, referred to as O∩R, representing the regions that both age groups recruited while performing the task; and an ROI consisting of the combination of both regions, referred to as O>R & O∩R. Contrasts were thresholded at $p < .05$ with whole-brain voxel-wise FWE correction, binarised, and inverse-normalised into participant native space. Within each ROI we fitted models using four contrasts that reflect different ways in which the stimuli may be represented: incongruence with 5° eccentricity (Incong5 ≠ Cong5), incongruence with 15° eccentricity (Incong15 ≠ Cong15), auditory location (AudL ≠ AudR), and visual location (VisL ≠ VisR). The outcome variable of interest for our analysis from each of these models is the log odds ratio, as it describes the degree to which the activity in a specific region can predict the contrast of interest. This is specified as the difference between the log evidence for the model of interest and a null model. We compared MVB models within and between groups in the following ways. First, we used Bayesian model selection (BMS; Rigoux, Stephan, Friston, & Daunizaeu, 2014) to assess which of the three ROIs held the most information about the stimuli in each age group. BMS assigns each model a protected exceedance probability that quantifies the probability of that model winning more frequently than other models across participants. We then compared the log odds ratio between age groups for each model, using both standard *t*-tests and a Bayesian equivalent, to assess whether any model fit better for one age group than another.

Support vector regression

Finally, we used a machine-learning approach to investigate whether age affects how congruent and conflicting audiovisual stimuli are represented in a-priori defined regions across the auditory and visual cortical hierarchies. In each ROI, we trained a linear SVR machine (using LIBSVM as implemented in The Decoding Toolbox Version 3; Hebart, Görgen, & Haynes, 2015) on the mapping between voxel response patterns and true stimulus location. The model was trained on labelled parameter estimate images of congruent stimuli (i.e. AV = -15, AV = -5, AV = 5, and AV = 15) from 10 of the 11 experimental runs. It was then given unlabelled parameter estimate images from congruent stimuli in the final run as test data and produced predicted labels for each image. Using a *k*-fold leave-one-run-out cross validation scheme, we repeated this procedure $k = 11$ times, leaving a different run out for testing each time. The per-location means of the predicted values across these 11 runs were taken as the final predicted values in each ROI. We then repeated this procedure, again training the machine on congruent stimuli, but instead used parameter estimate images of incongruent stimuli (i.e. A = -15 V = 15; A = -5 V = 5; A = 5 V = -5; and A = 15 V = -15) as the test data. In this way, we were able to distinguish which features of the congruent stimuli contributed most to the representation in each cortical region. For example, if a model trained on congruent data gave labels that matched the visual location when shown incongruent data, this would imply that the region in question primarily encoded information about the visual component of the stimulus.

This analysis was conducted individually for each participant in native space. Visual (V1 – V3) and intraparietal sulcus (IPS 0 – 2, IPS 3 – 4) ROIs were defined using retinotopic maximum probability maps (Wang, Mruczek, Arcaro, & Kastner, 2015); primary auditory cortex (A1) was defined based on cytoarchitectonic probability maps (Eickhoff et al., 2005), each inverse-normalised to participants' native space. As planum temporale (PT) is variable in size and may be defined anatomically (Shapleske, Rossell, Woodruff, & David, 1999), we used Freesurfer's automatic Destrieux parcellation (Fischl, 2012) to define this ROI in each participant individually.

For statistical analysis the SVR predicted labels were treated in a similar way to the behavioural data: separately for each ROI, participants' mean responses were entered into a 2 (congruent/incongruent) x 4 (location) x 2 (age) mixed ANOVA. For analyses and plotting, locations in the incongruent condition were labelled based on the dominant sense in each ROI: data from V1 – V3, IPS 0-2, and IPS 3-4 were labelled in terms of visual stimulus location, while data from PT and A1 were labelled by auditory location. The benefit of this approach is that it allows the congruence x location interaction to directly index the impact of the non-dominant modality on predicted location, and also allows for easy visual comparison between congruent and incongruent conditions when plotted (see Figure 4.5).

Results

Behavioural results

In a behavioural screening session, participants completed a left-right auditory localisation task to assess their ability to accurately localise stimuli with simulated location cues. There was no significant difference between age groups for either JND (accuracy), *t*(30)

= 1.532, $p = .136$, or PSE (left/right bias), $t(30) = 0.527$, $p = .602$, values. Participants also took part in two runs of the same behavioural task as was used during scanning. Group-level correlation coefficients show that participants were able to accurately locate the sound under unimodal (younger $r = .972$, older $r = .955$) and congruent (younger $r = .985$, older $r = .966$) conditions, and were far worse under incongruent conditions (younger $r = .622$, older $r = .266$). A repeated-measures ANOVA confirmed this with a strong main effect of location, $F(1.410, 42.304) = 310.537$, $p < .001$, $\eta^2_p = .912$, and a congruence x location interaction, $F(1.574, 47.214) = 33.975$, $p < .001$, $\eta^2_p = .531$. No other main effects or interactions were significant at $p < .05$.

We then investigated how audiovisual spatial congruence or incongruence affected participants' auditory localisation responses inside the scanner. Figure 4.1 shows group mean response for each location under each condition. Perfect performance would fall on the $y = x$ diagonal. As confirmed by group-level correlation coefficients, participants again performed well under unimodal conditions (younger $r = .948$, older $r = .899$) better under congruent audiovisual (multisensory enhancement; younger $r = .970$, older $r = .958$), and much more poorly under incongruent conditions (ventriloquist effect; younger $r = .571$, older $r = .172$). A mixed-measures ANOVA found a strong main effect of location, $F(1.239, 37.157) = 203.280$, $p < .001$, $\eta^2_p = 0.871$, and a congruence x location interaction, $F(1.364, 40.926) = 40.068$, $p < .001$, $\eta^2_p = 0.572$. No other main effects or interactions, including any involving age, were significant at $p > .05$.

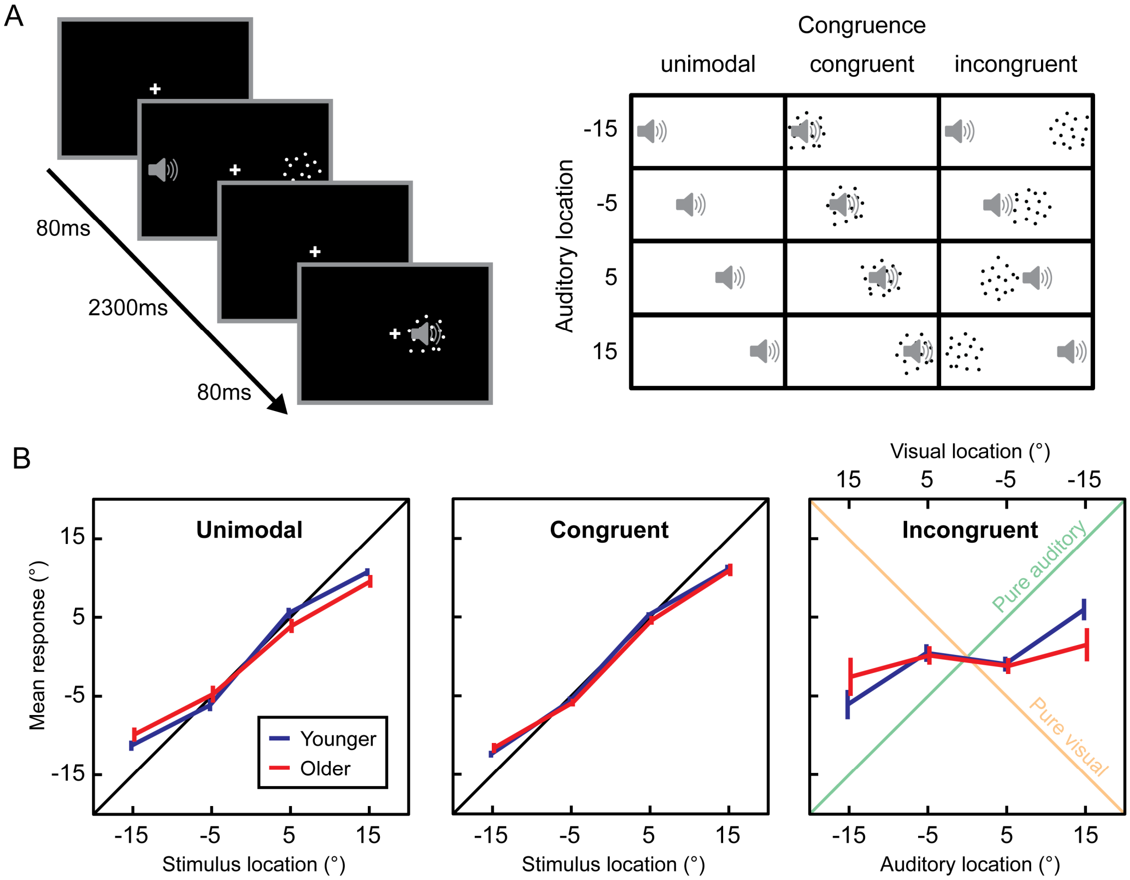

Figure 4.1. (A) Two example trials and factorial experimental design. In a symmetrical spatial ventriloquist paradigm, participants were presented with auditory stimuli (white noise) sampled from four possible azimuthal locations (-15°, -5°, 5°, or 15°). In one third of trials, the sound was presented alone. In the remaining two thirds, the sound was accompanied by a simultaneous visual (cloud of dots) stimulus that could be either spatially congruent or symmetrically incongruent. The design could therefore be characterised as a 4 (auditory location) x 3 (congruence) factorial paradigm. Participants reported the perceived location of the sound. (B) Behavioural responses in unimodal, congruent and incongruent conditions. Older participants appear marginally worse than younger adults at localising unimodal auditory stimuli, but this small difference disappeared for congruent stimuli: participants in both age groups were similarly able to localise congruent audiovisual stimuli and displayed a similar central bias (indicated by deviation from the diagonal at the extreme locations). For spatially incongruent stimuli, the magnitude of the ventriloquist effect (indicated by a change in slope away from the pure-auditory line) was near-identical between groups at small spatial disparities. At larger disparities, older adults showed a slightly stronger effect. No group differences were significant at $p < .05$, however.

fMRI results

General linear model

To confirm the effectiveness of the experimental manipulation we first checked for an overall effect of stimulus incongruence across groups. A network of frontal and parietal regions showed significantly greater activation for incongruent compared to congruent stimuli. Incongruence effects were apparent in superior parietal regions including right precuneus. We also observed activations bilaterally in superior frontal sulcus, in supplementary motor area on the medial surface of left frontal lobe, and bilaterally in insula. A summary of these results can be seen in Figure 4.2.

We then tested for areas where older adults showed significantly more activation than younger participants overall when performing the task. This revealed a number of frontal, parietal, and superior temporal regions, though the profile of age differences across these areas varied. Superior frontal activations, including some likely to be associated with motor responses, were apparent in both groups, but the older adults showed higher overall activation. Conversely, regions of intraparietal sulcus (IPS) bilaterally showed deactivation in younger adults, but positive responses in the older adults. Other areas, including left planum temporale and inferior frontal gyrus, showed very limited activity in the younger age group but substantial responses in the older adults. See Figure 4.3 for an overview. The equivalent test for activations that were greater in younger than older adults produced no significant results. To formally test for age differences in congruence-related activity we also checked for age x congruence interactions, but none were significant after correction for multiple comparisons.

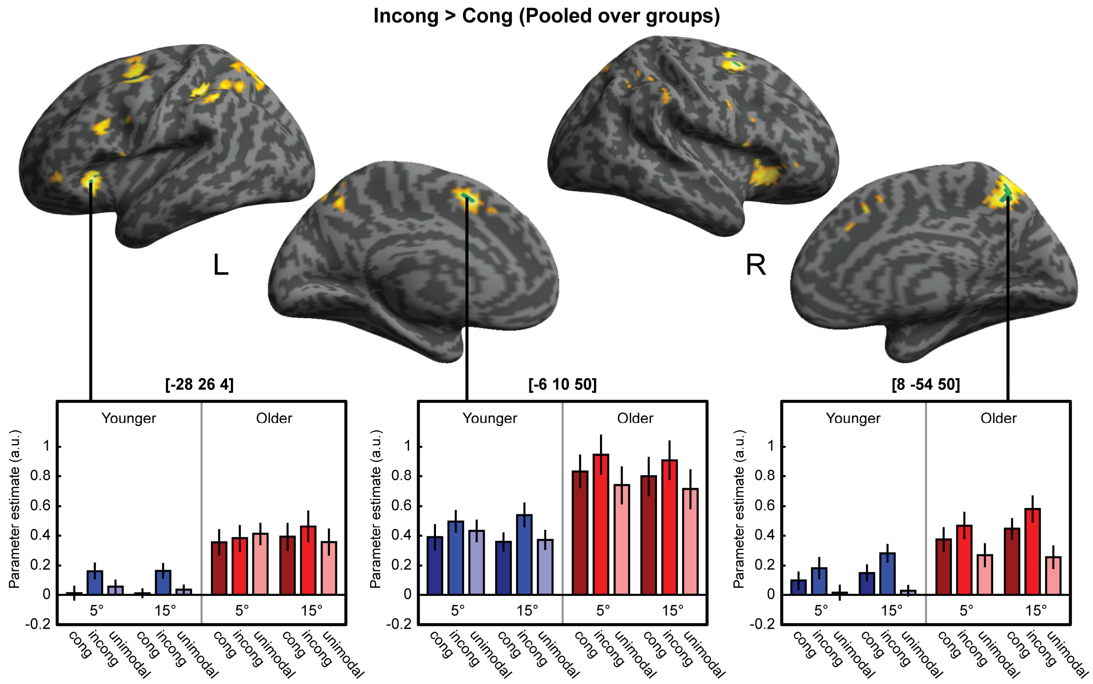

Figure 4.2. fMRI results showing activation increases for congruent relative to incongruent conditions, pooled over age groups, rendered on an inflated template brain. Results are thresholded at $p < .001$ for illustration; voxels that survive whole-brain FWE correction at $p < .05$ are highlighted in green. Parameter estimates, averaged over left/right stimuli, are plotted for peaks at the MNI coordinates indicated. Error bars show ±1 *SEM*.

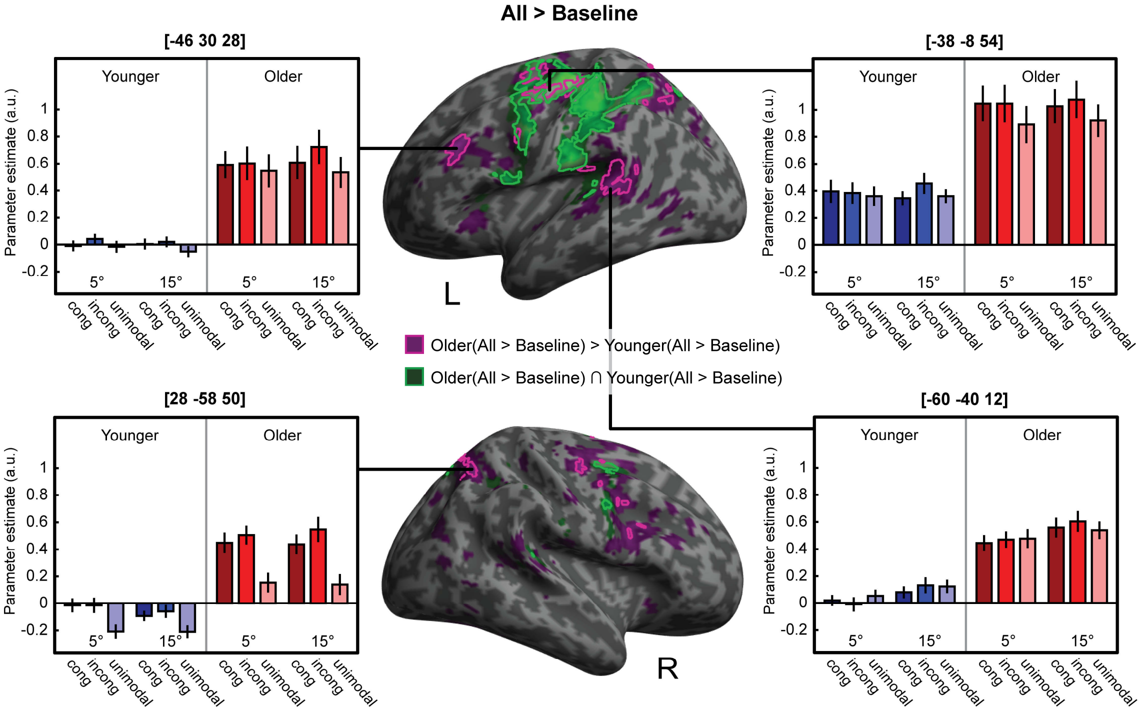

Figure 4.3. fMRI results showing between-group similarities and differences in task-related activation versus baseline, rendered on an inflated template brain. The outcome of the (All_{Older} > $Baseline_{Older}$) > ($All_{Younger}$ > $Baseline_{Younger}$) interaction contrast is shown in purple; the conjunction analysis (All_{Older} > $Baseline_{Older}$) ∩ ($All_{Younger}$ > $Baseline_{Younger}$) is in green. Darker colours show results thresholded at $p < .001$ uncorrected, while the brighter edges indicate voxels that survive $p < .05$ whole-brain FWE correction. Parameter estimates, averaged over left/right stimuli, are shown for peaks at the MNI coordinates indicated. Error bars show ±1 *SEM*.

Table 4.1.

fMRI results: Group-level GLM

Region	Coordinates			*z*-score	*p*-value (FWE*)
Incong > Cong (Pooled over age groups)					
R. precuneus	8	-54	50	5.2	< .001
L. supplementary motor area	-6	10	50	5.0	< .001
L. superior frontal sulcus	-26	6	58	5.0	< .001
L. superior frontal sulcus	-26	-2	48	4.9	< .001
L. anterior insula	-28	26	4	5.0	< .001
R. superior frontal sulcus	24	2	54	4.8	< .001
R. anterior insula	32	26	-4	4.8	< .001
L. superior frontal sulcus	-30	-2	62	4.7	< .001
O(All > Baseline) > Y(All > Baseline)					
L. inferior frontal sulcus	-46	30	28	7.3	< .001
L. precentral gyrus	-38	-8	54	6.6	< .001
L. supplementary motor area	-8	-8	64	6.3	< .001
L. superior frontal sulcus	-20	-8	56	5.8	< .001
L. superior temporal gyrus	-60	-40	12	5.8	< .001
L. superior temporal gyrus	-46	-34	16	5.6	.001
L. supramarginal gyrus	-50	-44	22	5.4	.001
R. intraparietal sulcus	28	-58	50	5.6	.001
R. precuneus	12	-62	62	5.5	.001
R. intraparietal sulcus	24	-62	56	5.0	.011
R. precentral sulcus	48	-4	52	5.6	.001
R. supplementary motor area	8	18	46	5.5	.001
R. inferior frontal sulcus	36	2	36	5.4	.002
L. precuneus	-10	-64	58	5.3	.002
L. intraparietal sulcus	-26	-70	50	5.2	.004
R. superior frontal sulcus	26	-6	56	5.2	.004
R. supplementary motor area	10	6	56	5.2	.005
R. superior frontal sulcus	26	6	54	5.2	.005
L. precentral sulcus	-46	6	34	5.1	.007
L. precentral sulcus	-50	-8	46	5.0	.012
L. intraparietal sulcus	-28	-54	46	4.9	.014
L. superior temporal pole	-52	14	-4	4.9	.018
R. inferior frontal sulcus	38	14	26	4.9	.019

L. intraparietal sulcus	-24	-62	58	4.8	.031
L. intraparietal sulcus	-44	-40	34	4.7	.037
L. anterior insula	-30	24	0	4.7	.047
O(All > Baseline) ∩ Y(All > Baseline)					
R. cerebellum	22	-54	-24	> 8	< .001
R. cerebellum	6	-62	-16	6.9	< .001
R. cerebellum	8	-72	-16	5.9	< .001
L. precentral gyrus	-36	-20	64	> 8	< .001
L. precentral sulcus	-32	-4	58	> 8	< .001
L. intraparietal sulcus	-46	-34	42	> 8	< .001
L. supplementary motor area	-4	0	56	> 8	< .001
R. superior frontal sulcus	24	-2	50	5.7	< .001
L. thalamus	-14	-18	6	5.4	0.002
L. intraparietal sulcus	-18	-68	54	5.4	0.002
R. precentral gyrus	52	4	42	5.1	0.005
L. planum temporale	-40	-36	10	5.1	0.007
L. anterior insula	-30	18	8	5.0	0.009
L. superior frontal gyrus	-16	-6	68	5.0	0.011
R. intraparietal sulcus	14	-66	52	4.9	0.014
R. superior temporal gyrus	58	-34	14	4.8	0.027

**p*-values whole-brain corrected for familywise errors at the voxel level

Multivariate Bayesian decoding

To better understand the age differences in regional activations revealed by the mass-univariate analysis we fitted a series of multivariate Bayesian decoding models. Models were based on data from three functionally-defined ROIs: the areas that older adults activated more than younger adults when performing the audiovisual localisation task (O>Y), the areas that both age groups activated when performing the task (O∩Y), and a combination of both of these as a full model (O>Y & O∩Y). See the bright outlines in Figure 4.3 for the location of these ROIs. Data from these models attempted to separately predict four different target variables defined by GLM contrasts: visual left versus visual right, auditory left versus auditory right, incongruent versus congruent at 5° eccentricity, and incongruent versus congruent at 15° eccentricity. Quality of fit was characterised in terms of the difference in log evidence between the model of interest and a null model. These fits were then compared within each target variable using Bayesian model selection to assess which best predicted the data, and between groups using both standard *t*-tests and a Bayesian equivalent. If older adults were using the areas where they show significantly greater activation in a compensatory fashion we would expect some of the following outcomes for one or more of the target variables: better model fit in older adults for the O>Y ROI; worse model fit in older adults for the O∩Y ROI; and/or greater benefit to older than younger adults for the full model (O>Y & O∩Y) versus the O∩Y model alone.

The results, however, strongly indicate that each of the ROIs hold similar amounts of information in the older and younger adults. First, Bayesian model selection indicated that the full model (O>Y & O∩Y) won for each of the target variables in all participants (exceedance probability of 1 in each group), suggesting that both age groups recruit regions from each of the ROIs to successfully perform the task. Second, pairwise comparisons of all log odds ratio values for each target variable in each model revealed no significant differences between age groups. See Table 3.1 for a summary. This suggests that older adults have similarly informative representations of the stimuli as younger adults in the O∩Y ROI, and that they do not hold more information about the stimuli in the O>Y ROI. The finding that the model fit for (O>Y & O∩Y) was equally good between age groups is also in line with our behavioural results, which show little age-related difference in task performance. These results are summarised in Figure 4.4.

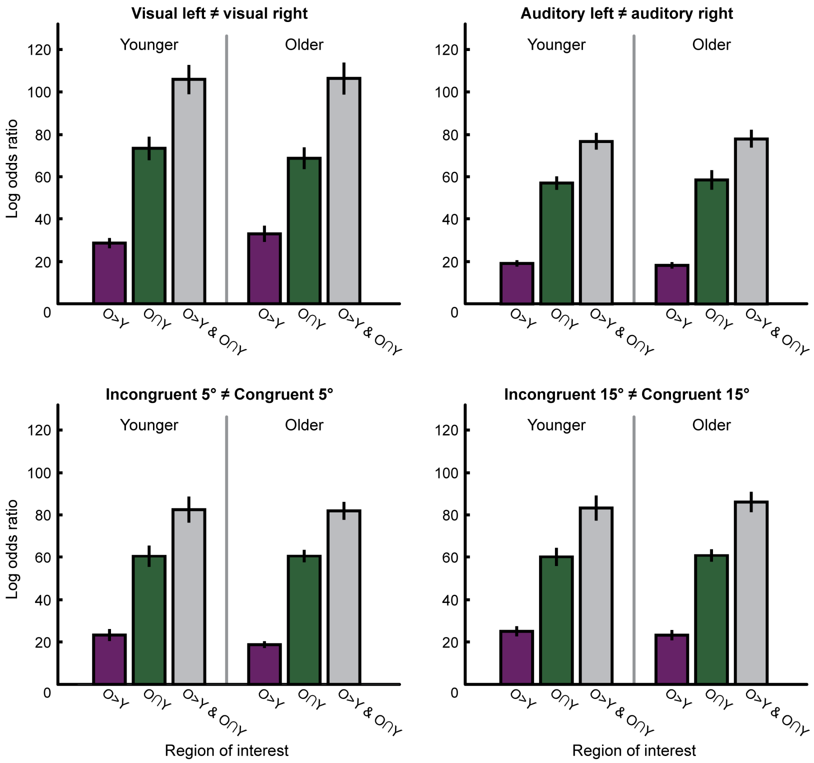

Figure 4.4. Results of multivariate Bayesian decoding analysis (MVB). Three models, based on activation data in functionally-defined regions of interest, were fitted for four stimulus contrasts. Bars show group means of log odds ratios, defined as the difference between the log model evidence of the fitted model and a null model. Error bars show ±1 *SEM*.

Table 4.2.

Comparison between age groups of log odds ratios for fitted MVB models

Target variable and ROI	BF_{10}	BF_{01}	*t*	*p*	Cohen's *d*
Visual left ≠ right					
O>Y	0.482	2.077	0.969	0.340	0.343
O∩Y	0.390	2.566	-0.620	0.540	-0.219
O>Y & O∩Y	0.337	2.971	0.048	0.962	0.017
Auditory left ≠ right					
O>Y	0.361	2.769	-0.431	0.670	-0.152
O∩Y	0.345	2.896	0.263	0.794	0.093
O>Y & O∩Y	0.342	2.928	0.202	0.841	0.071
Congruent 5° ≠ Incongruent 5°					
O>Y	0.699	1.430	-1.391	0.174	-0.492
O∩Y	0.336	2.974	0.008	0.994	0.003
O>Y & O∩Y	0.337	2.967	-0.077	0.939	-0.027
Congruent 15° ≠ Incongruent 15°					
O>Y	0.376	2.660	-0.539	0.594	-0.190
O∩Y	0.338	2.955	0.128	0.899	0.045
O>Y & O∩Y	0.355	2.819	0.373	0.712	0.132

Note. Student's two-tailed independent-samples *t*-test and Bayesian equivalent. BF_{01} quantifies degree of support for the null hypothesis that there is no difference between groups; BF_{01} shows degree of support for the alternative hypothesis that the groups differ.

Support-vector regression

As we have a-priori knowledge about the cortical regions involved in audiovisual spatial integration, we also aimed to evaluate whether age influences the patterns of activity in these areas. In various ROIs across the cortical hierarchy, support vector regression models were trained on BOLD responses to congruent stimuli and used to predict labels for congruent and incongruent data. We entered the predicted locations into separate 2 (congruence) x 4 (stimulus location) x 2 (age) mixed ANOVAS. The profiles of decoded locations fit well with previous research (Rohe & Noppeney, 2015, 2016). Importantly, a main effect of stimulus location was present for all ROIs at $p < .001$, suggesting all regions encoded some information about the location of the stimuli. See Figure 4.5 for a summary of these results. As anticipated, areas V1 – V3 were found to encode primarily the location of the visual stimulus. However, a congruence x location interaction was also observed, $F(3, 90) = 10.929$, $p < .001$, $\eta^2_p = .267$, indicating that the location of the auditory stimulus had some influence on representations even in lower-level visual areas. Posterior intraparietal sulcus (IPS 0-2) showed a similar profile to V1 – V3, with a small but highly-significant effect of stimulus congruence (i.e. sound location), $F(2.137, 64.104) = 6.302$, $p = .003$, $\eta^2_p = .174$. Decoded locations in anterior IPS (3 – 4) resemble ventriloquist behavioural responses, with visual location having a stronger influence at the small disparity level; the congruence x location interaction was thus more pronounced here, $F(3, 90) = 11.147$, $p < .001$, $\eta^2_p = .271$. Planum temporale also showed a strong congruence x location interaction, $F(2.322, 96.659)$, = 7.255, $p < .001$, $\eta^2_p = .195$, again with visual stimuli having a greater impact at smaller disparity levels. Primary auditory cortex represented purely the location of the auditory stimuli and did not show a significant congruence x location interaction, $F(3, 90) = 2.536$, $p = .062$, $\eta^2_p = .078$. Crucially, none of the ROIs showed a main effect of, or interactions with, age group, p

> .05. It is apparent in Figure 4.5 that representations of stimulus locations were remarkably similar between age groups across all ROIs and for both congruent and incongruent stimuli.

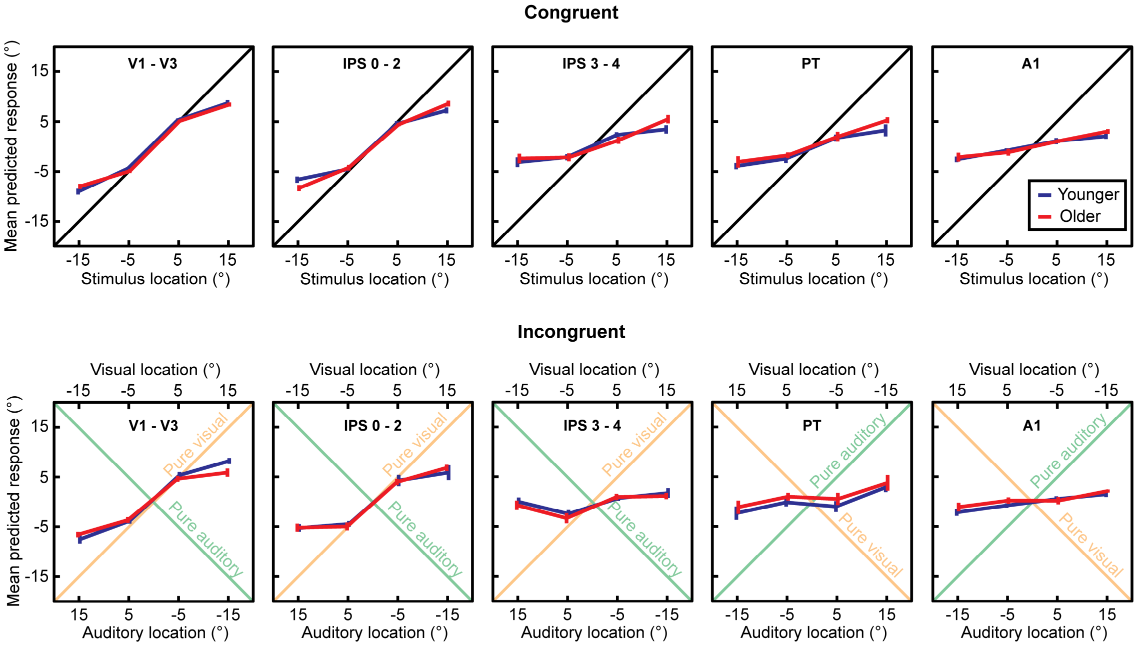

Figure 4.5. Results of support vector regression (SVR) multivariate pattern analysis (MVPA). A support vector machine was trained on BOLD responses to spatially congruent audiovisual stimuli for various cortical ROIs separately. The top row shows the predicted labels when the trained model was provided BOLD responses to novel congruent stimuli as test data. The bottom row shows predicted labels when the same trained model was tested on BOLD responses to symmetrically incongruent stimuli. Results for five ROIs are shown: visual regions (V1 - V3); posterior intraparietal sulcus (IPS 0 - 2); anterior intraparietal sulcus (IPS 3 - 4); planum temporale (PT); and primary auditory cortex (A1). Lines show group averages; error bars show ±1 *SEM*. Note that in plots of incongruent results the location of the auditory stimulus is plotted on the lower x-axis and the visual on the upper x-axis, and that the direction of these axes is dependent on the ROI's dominant sensory modality. The green and orange lines indicate the expected response if the predicted values were based entirely on auditory or visual stimulus location respectively.

Discussion

This study investigated the neural mechanisms underlying integration of multisensory spatial information in younger and older adults. Participants viewed spatially congruent and incongruent audiovisual stimuli and reported the location of the sound while undergoing functional MRI scanning. In agreement with previous research, the groups gave similar behavioural responses. Both younger and older participants were able to accurately localise sounds inside the scanner, and these responses improved when a spatially congruent visual stimulus was also presented. Accuracy declined significantly in the presence of a spatially incongruent distractor, especially when this was presented near to the location of the sound, indicating a robust ventriloquist effect. Although it did not reach statistical significance, there was a trend towards reduced accuracy in the older group that should be acknowledged. Occurring particularly in the unimodal condition and for incongruent stimuli at the larger spatial disparity, this effect is consistent with slightly reduced auditory reliability in the older adults, probably due to scanner noise having a disproportionately negative effect on this age group: older adults are known to struggle more with interference from background noise (Alain, McDonald, & Van Roon, 2012; Dubno, Dirks, & Morgan, 1984; Presacco, Simon, & Anderson, 2016).

Standard mass-univariate GLM analyses conducted on both age groups together revealed a network of regions that responded more to spatially incongruent than congruent stimuli. As well as activations that are probably associated with response planning or execution, such as those in supplementary motor area, we also found incongruence effects in superior parietal lobule, which is consistent with previous work on audiovisual spatial integration (Avillac, Ben Hamed, & Duhamel, 2007; Macaluso & Driver, 2005; Rohe & Noppeney, 2015b, 2016; Schlack et al., 2005). Activity was also apparently bilaterally in anterior insula (but only survived whole-brain multiple-comparison correction in the left hemisphere), a critical part of the salience network and an area regularly activated in tasks that require response suppression to an incongruent stimulus (Menon & Uddin, 2010; Roberts & Hall, 2008). We also saw activation in superior frontal sulcus, which has previously been implicated in spatial working memory (Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998)

We then used the GLM analysis to explore age-related activation differences. Age did not significantly interact with any stimulus condition, but a between-group comparison in overall task activation revealed an array of frontal, parietal, and superior temporal regions that were activated more in older adults. This fits closely with previous work showing older adults recruiting extra regions during a variety of tasks (Davis, Dennis, Daselaar, Fleck, & Cabeza, 2008; Reuter-Lorenz & Cappell, 2008; Morcom, Henson, & Cam-CAN, 2018), including those that involve processing of complex multisensory stimuli (Townsend, Adamo, & Haist, 2006). The diversity in activation profiles across these various regions implies that there may not be one single mechanism driving the age differences, however. For example, left motor regions showed large significant activation in both age groups, as expected, and only the magnitude varied between groups. Conversely, older adults had strong activations for all conditions in left inferior frontal gyrus (IFG), while the younger age group barely recruited this area. Left IFG has been shown to be critical for response inhibition (Swick, Ashley, & Turken, 2008), and it is thus possible that older adults required extra input from this region to inhibit responses to visual stimuli. Regions of right intraparietal sulcus, discussed in further detail below, deactivated in response to unimodal auditory stimuli in younger adults, but showed positive responses to all stimulus types in the older age group.

However, MVB decoding analyses revealed that the quantity of stimulus-related information in these regions did not differ between age groups: all areas that activated more in older adults held equal levels of information in younger participants. Furthermore, regions that were significantly activated in both groups also encoded similar information between age groups, suggesting that there was no degradation here to be compensated for. This speaks against theories of age-related shifts in activation as compensatory mechanisms (Cabeza, 2002; Davis et al., 2008; Park & Reuter-Lorenz, 2009; Reuter-Lorenz & Cappell, 2008), but is supported by recent work that applied a similar MVB decoding approach to long-term memory encoding and short-term memory maintenance tasks in large samples of younger and older adults (Morcom et al., 2018). The authors observed significantly increased prefrontal activations in older adults for both tasks, but found that this region encoded comparable levels of information in both age groups.

The argument against age-related activation differences in the present study being compensatory is further supported by our findings that the established cortical hierarchy for audiovisual integration is intact in older adults. Although our simplified ventriloquist paradigm (which was chosen primarily to be shorter and thus more tolerable for older adults) is unsuitable for the computational analyses applied by Rohe and Noppeney (2015, 2016), the results show the same pattern. Low-level sensory regions encode the location of their respective modality, while anterior parts of intraparietal sulcus encode the final fused (or segregated) percept. Our analysis was indeed sufficiently sensitive to detect effects of stimulus incongruence even in lower-level visual regions. As the representations in these regions are not degraded by age, the necessity for compensatory activity in other regions is called into question.

It is important to acknowledge the possible influence of motor responses in the SVR decoding analysis. As noted, inclusion of the button response as a regressor in the GLM reduced the accuracy of this decoding, but this does not necessarily mean that the activity being decoded is primarily motor-related: in this paradigm, the response is entirely confounded by perceived location of the sound (participants pressed to indicate where they perceived the sound to be), so the two cannot be separated. It is, nonetheless, a possibility that must be considered when interpreting the results. We do not believe that this is a significant limitation of the design, however, as our primary outcome of interest was age differences in representation and not the nature of the representations per se. Furthermore, our conclusions are supported by analyses that are orthogonal to the button responses (MVB decoding of congruence).

We also acknowledge that the findings could be enriched by further, response-driven analysis. Some analyses of this type are challenging with the present design. As noted above, our symmetrical paradigm precludes the computational comparisons of behavioural and neural data performed by e.g. Rohe and Noppeney (2015, 2016), as it confounds the ventriloquist effect with the spatial prior (the ventriloquist effect is always towards the centre of space in our design). Similarly, the application of machine learning approaches such as support vector regression becomes substantially more complex with unbalanced data (as is the case with response-driven regressors). It may, however, be possible to apply new (response-driven) GLMs to the present dataset to further explore age differences under specific conditions. For example, some differences in the processing of incongruent stimuli might only be present on trials where a ventriloquist effect occurs; these would be missed by our current GLM analysis, but could be revealed by separating such trials into a separate regressor. Future re-analysis of these data may thus reveal more nuanced age differences in activation amplitudes than are reported here.

In conclusion, this study showed that older and younger adults utilise the same network of regions for audiovisual spatial integration, but that some areas are activated more in older adults. This activity may reflect simple inefficiency, as argued by Morcom et al. (2018), but is equally likely to be due to increased effort on the part of older adults, perhaps in response to finding auditory localisation in a noisy environment more challenging.

CHAPTER 5: HEALTHY AGEING AFFECTS AUDIOVISUAL SPEECH COMPREHENSION IN CONDITIONS OF ASYNCHRONY AND BACKGROUND NOISE: A BEHAVIOURAL AND fMRI STUDY

Samuel A. Jones, Giulio Degano, and Uta Noppeney

Difficulty understanding speech in a noisy environment is a common complaint in older adults, and frequently reproduced in the laboratory (Anderson Gosselin & Gagné, 2011; Dubno, Dirks, & Morgan, 1984; Helfer & Freyman, 2008; Wilson, McArdle, Watts, & Smith, 2012). Some of the effect is attributable directly to presbycusis (Humes, 2002, 2007), but impaired comprehension of speech in noise has been demonstrated even in healthy older adults with clinically normal pure-tone hearing thresholds, leading researchers to conclude that central auditory and cognitive factors also play a role (Füllgrabe et al., 2015; Helfer & Freyman, 2008; Humes, 2007).

The addition of congruent visual information can help both younger and older adults to detect and comprehend auditory speech signals that are embedded in background noise; some studies have found that the degree of visual enhancement is similar between age groups (Anderson Gosselin & Gagné, 2011; Helfer, 1998; Tye-Murray, Sommers, Spehar, Myerson, and Hale, 2010; Tye-Murray, Spehar, Myerson, Hale, and Sommers, 2016), while others report that older adults receive less benefit (Dey and Sommers, 2015; Stevenson et al., 2015). In younger adults, the cortical mechanisms underlying audiovisual speech perception have been investigated extensively. Low-level multisensory interactions in primary auditory cortex notwithstanding (Ghazanfar, Maier, Hoffman, and Logothetis, 2005; Okada, Venezia, Matchin, Saberi, and Hickok, 2013; Pekkola et al., 2005), integration of audiovisual speech information appears to take place principally in superior temporal sulcus (STS) and superior temporal gyrus (STG), collectively referred to as superior temporal cortex (STC). Multisensory superior temporal cortex (mSTC) has previously been functionally defined as the area of superior temporal lobe that responds to both auditory and visual speech stimuli (Stevenson, Altieri, Kim, Pisoni, & James, 2010; Stevenson, VanDerKlok, Pisoni, & James, 2011). Various functions are spread along the length of mSTC, with anterior regions representing primarily auditory information and posterior areas responding more to visual and integrated stimuli, as well as representing synchrony (Stevenson et al., 2010; Stevenson et al., 2011; Ozker et al., 2017). However, these mechanisms have received limited attention in the older population, despite evidence that ageing leads to changes in patterns of neural activity in response to multisensory signals (Setti, Finnigan et al., 2011; Stephen, Knoefel, Adair, Hart, & Aine, 2010; Townsend, Adamo, & Haist, 2006).

Furthermore, the perception of multisensory stimuli has been shown to change throughout the human lifespan, particularly when there is incongruence in the temporal structure of the signals. Audiovisual synchrony perception develops in infancy (Lewkowicz, 1996, 2010), and declines in later life (Bedard and Barnett-Cowan, 2015; Colonius & Diederich, 2011; McGovern, Roudaia, Stapleton, McGinnity, & Newell, 2014; Setti, Burke, Kenny, & Newell, 2011). This reduced temporal sensitivity may partially account for older adults' impaired multisensory speech perception, including greater susceptibility to audiovisual speech illusions (Sekiyama, Soshi, & Sakamoto, 2014; Setti, Burke, et al., 2013), but investigations into this possibility are limited. Stevenson, Baum, Krueger, Newhouse, & Wallace (2017) showed, across a large range group of participants aged between five and eighty years old, that age affected both temporal acuity and susceptibility to the McGurk illusion, but these did not correlate at the individual level. Conversely, Başkent & Bazo (2011) found that asynchrony detection and speech comprehension abilities did correlate in their sample of healthy younger and hearing-impaired older adults. Further work that directly investigates the relationship between temporal acuity and speech impairment in older adults is clearly needed.

The present study aimed to investigate younger and older adults' behavioural and neural responses to multisensory speech stimuli presented under challenging circumstances, and to explore the complex interplay between age, background noise, asynchrony, and audiovisual speech comprehension. Specifically, we engage with the following questions: Does a congruent visual signal improve comprehension of speech in noise to a similar degree in younger and older adults? Can audiovisual asynchrony disrupt comprehension of speech in noise, and is this effect stronger in older adults? Does the presence of background noise impair detection of asynchrony? Is the ability to detect asynchrony related to its effect on comprehension, and does this change with age? Finally, how does ageing affect neural responses to audiovisual speech stimuli presented under these various conditions?

In a behavioural testing session, we tested younger and older participants' responses to unimodal, synchronous, and asynchronous spoken sentence stimuli, both in a quiet environment and with disruptive background noise. First, we assessed participants' ability to detect asynchrony in the stimuli. We then employed a dual-response paradigm with a 2 (age) x 5 (synchrony) x 2 (background noise) factorial design to explicitly and implicitly measure speech comprehension under these various conditions. In a later session, participants repeated the explicit comprehension task while undergoing fMRI scanning.

Methods

Participants

Fifteen younger (7 female, mean age = 21.87, *SD* = 3.40, range = 19 – 30 years) and 15 healthy older (8 female, mean age = 70.87, *SD* = 4.76, range = 67 – 83 years) native speakers of British English were recruited from local participant databases to take part in the study. All participants had normal or corrected-to-normal vision. Older adults showed mild to moderate hearing loss consistent with that expected of adults in their age group (Wiley, Chappell, Carmichael, Nondahl, & Cruickshanks, 2008). See Figure 5.2 for a summary. Older adults were screened for cognitive impairment using the Montreal Cognitive Assessment (Nasreddine et al., 2005); none scored below 25. All participants gave informed consent to participate. The study was given a favourable ethical opinion by the University of Birmingham Ethical Review Committee.

Stimuli

The stimuli were 416 sentences (58%) taken from the 720 IEEE/Harvard (1969) sentences. These were taken from the beginning of the list, though we removed any that we judged might be difficult for a modern speaker of British English to understand (as the list was created around 50 years ago in the US, some words and expressions may be unfamiliar to our participants; e.g. *The cold drizzle will halt the bond drive*). Sentences were assigned five target words, allowing participants' verbal repetition of each sentence to be scored out of five. Target words in the first sentence, for example, were: *The* <u>*birch*</u> <u>*canoe*</u> <u>*slid*</u> *on the* <u>*smooth*</u> <u>*planks*</u>. The sentences were spoken by a female British English speaker against a dark-blue background in a sound-isolated chamber. Video was filmed at a high sampling rate of 60 frames per second on a Samsung Galaxy S7 smartphone. Audio was recorded through a Shure

SM57 dynamic microphone, via an Alesis iO2 Express external soundcard, onto a Macbook Pro sampling at 44100 Hz.

After these files were transferred to a PC, the following steps were conducted variously in Audacity (Audacity Team, 2018), PRAAT (Boersma & Weenink, 2018), Shotcut (Meltytech, 2018), and MATLAB. Recorded video files were cropped to a resolution of 500 x 300 pixels to show only the lower half of the face with the lips centred in the frame. Audio files were cleaned to remove background noise. Video and audio data were then segmented into separate files for each sentence, based upon manual demarcation of sentence onset and offset in the audio track, and the auditory files were normalised for amplitude. The final video files began 300ms before the onset of each sentence, and finished 300ms after the end of the last word, to allow stimuli to begin and end with a closed mouth. Auditory files only contained sound information from between the marked onset and offset and were padded with silence to match the duration of the video. Separate auditory files were created for every sentence at various asynchrony levels, with the onset of the sound shifted relative to the video, allowing all stimuli to be presented under any experimental condition. We refer to the asynchrony levels in terms of the temporal position of the sound relative to the video, so an asynchrony level of -250 indicates that the auditory track began 250 ms prior to the video.

We ensured the synchrony of the auditory and video files by filming a custom LED/beeper circuit at the beginning and end of each recording session. This circuit, described in detail by Maier, Di Luca, and Noppeney (2010), presents highly synchronous 500ms LED flashes and 1kHZ beeps every 1000ms. The temporal alignment of the auditory and visual tracks was adjusted manually so that the onsets of these signals matched. Due to the high frame rate of the video recordings we were able to achieve a good level of synchrony in the final stimuli: the mean difference between the auditory and video onsets was 0ms with an *SD* of 7ms (measured by filming and presenting the synchrony circuit under the same conditions as the stimuli).

The babble noise was created by recording five female undergraduate students separately speaking 230 sentences (32%) from the remaining IEEE Harvard list that had not been used for the stimuli. The sentences were segmented into individual files and their amplitude normalised. A track was then created for each speaker by randomising the order of the sentences and concatenating them with minimal gaps. These tracks were mixed together with a small temporal offset and trimmed at the beginning and end to ensure all speakers were included throughout. Finally, a dynamic range compression filter was applied to stabilise the dynamic range (and therefore the speech-masking properties) of the signal.

Based on behavioural piloting, a speech/babble signal-to-noise ratio (SNR) of 0dB was selected for the noise condition as a level that interfered with sentence comprehension to a sufficient degree in both age groups to allow performance to be improved by a congruent visual signal, while still allowing for some comprehension under noise conditions. The SNR was calculated from the signals' root-mean-square (RMS) level according to the following formula.

$$SNR_{dB} = 20 \cdot log_{10}\left(\frac{RMS_{signal}}{RMS_{noise}}\right)$$

Behavioural session

At least two weeks prior to the fMRI scan, participants took part in a behavioural testing session consisting of an audiometric measurement, a synchrony judgement paradigm, and a dual-response sentence comprehension paradigm. During this session, each of the 416 sentences that would later be presented during the MRI scan was presented exactly once, randomised to task and condition individually for each participant. Visual stimuli were presented on an LCD monitor, positioned 55 cm from the participant, at a 60Hz refresh rate.

The speech videos were displayed in the centre of the screen at a width of 15°. Auditory stimuli were presented over Sennheiser HD 280 headphones, with the sentence recording played at approximately 75dB SPL. Stimuli were controlled via Psychtoolbox-3 (Kleiner, Brainard, & Pelli, 2007) running on MATLAB under Ubuntu 16.04.

Audiometry

Participants' hearing thresholds were evaluated in an automated audiometric procedure (Maltby, 2002). Beginning at 50 dB SPL, pure tones were presented for a randomised duration of between 1.5 and 4.5 seconds, and participants instructed to press and hold a button for the duration of the tone. If an accurate response was given, the amplitude of the tone decreased by 10dB SPL. If not, it increased by 5dB. This adaptive staircase continued until four reversals (volume decrease then increase) had been completed, and the procedure repeated for six frequencies (1000, 2000, 4000, 8000, 500, and 250 Hz) in each ear. The threshold for each frequency was taken as the lowest amplitude at which two reversals occurred. If no two reversals occurred at the same amplitude, an average was taken across all four reversals. Note that this audiometric testing was undertaken using a custom setup and without full sound isolation; absolute thresholds in both groups are slightly higher as a result.

Synchrony judgment

A synchrony judgement task was used to evaluate participants' ability to perceive audiovisual asynchrony in the sentence stimuli under quiet and noisy conditions. The task was constructed as a 5 (asynchrony levels of -250, -150, 0, 150, and 250 ms) x 2 (no noise/0 dB SNR) factorial paradigm, with 28 sentences presented in each condition. Stimuli were presented in two blocks, divided by noise level (the babble track was started before, and looped throughout, the noise block); the order of these blocks was counterbalanced between participants.

Audiovisual sentence stimuli were presented synchronously, or under one of the four asynchrony levels, in a randomised order with a jittered inter-trial interval of between 2.3 and 3 seconds. Participants were instructed to indicate with a button press whether they perceived the sentences to be synchronous or asynchronous. Between trials, participants fixated a cross centred at the location of the speaker's lips.

Dual-response sentence comprehension.

The behavioural task used in the MRI scanning session required participants to make a subjective evaluation of their ability to comprehend a sentence stimulus by rating their understanding on a scale (1 = "Understood nothing", 2 = "Understood some", and 3 = "Understood all"). It is possible that participants are inaccurate in such judgements, which would make behavioural responses given inside the scanner difficult to interpret. This dual-response paradigm therefore served to provide a more direct measure of each participant's ability to comprehend the stimuli under various asynchrony/noise conditions, and to assess the accuracy of their subjective responses.

The same five stimulus types were used for both this task and for the fMRI paradigm: synchronous audiovisual, asynchronous audiovisual (-250 ms), asynchronous audiovisual (+250 ms), unimodal auditory, and unimodal visual. Each type was presented under noisy and quiet conditions, resulting in a 5 x 2 factorial design. A total of 16 trials were included in each condition (8 in the asynchrony conditions), randomised and blocked by noise condition.

After each stimulus, participants were instructed to rate their understanding with a 3-alternative button press: 1 = "Understood nothing", 2 = "Understood some", 3 = "Understood all". Participants then repeated the sentence out loud as accurately as possible to the experimenter, who marked the number of correctly repeated words on a score sheet before cueing the next trial with a keypress. The volume of the babble noise (when present) was reduced during the response period so that participants could repeat the sentences without removing the headphones. Between trials and during auditory-only trials, participants were instructed to fixate a central cross. Participants were given a four-trial practice session with no babble noise to become familiar with task prior to the first block.

fMRI session

The MRI scanning session took place a minimum of two weeks later to reduce memory effects. In the scanner, auditory stimuli were presented through Optime 1 electrodynamic headphones (MR Confon). Visual stimuli were projected by a JVC DLA-SX21E projector onto an acrylic screen and viewed via a mirror attached to the MRI head coil. The total viewing distance from eye to screen was 68cm. Responses were made using infrared response pads (Nata Technologies) held in the right hand.

Immediately prior to entering the scanner, participants completed one run of the task (with babble noise) to re-familiarise them with the stimuli and responses. The stimuli used in this task were spare sentences that had not been included in the behavioural or fMRI paradigms. During the fMRI session, participants watched and listened to spoken sentence stimuli presented under A, V, AVsync, AV-250, and AV+250 conditions. Data were acquired in eight separate scanning runs; on every other run, the babble track was started at the beginning and played throughout. The order of run types was counterbalanced between participants. Including null trials, this resulted in a 6 (stimulus) x 2 (noise level) factorial design. Participants responded to each stimulus by rating their comprehension via a right-handed button press in the same way as in the behavioural session. A total of 52 trials was presented in each condition, except for the asynchrony conditions which contained 26 each and are pooled over for imaging analysis. All sentences in the dataset were presented exactly once and their assignment to conditions was counterbalanced between participants.

fMRI data acquisition

A 3T Philips Achieva MRI scanner was used to acquire T1-weighted anatomical images and T2*-weighted BOLD EPI (gradient echo, SENSE = 2, TR = 8800 ms, TA = 2800 ms, TE = 40 ms, flip angle = 90°, FOV = 192×192x114mm², 38 slices, voxel size = 2.5×2.5×2.5 mm3 with a 0.5 mm interslice gap). A temporally-sparse EPI sequence was utilised: each 2800 ms volume acquisition was followed by a 6000 ms period of silence, during which the stimuli were presented (see Figure 5.1C). Stimulus onset was randomly jittered such that each sentence could occupy any portion of the silent period. Eight ten-minute runs of 67 volume images each were collected. We also acquired a five-minute resting state scan, using the same EPI sequence without the 6000 ms gap (TA = TR = 2800 ms), throughout which participants fixated a central cross.

fMRI data analysis

Functional fMRI data were analysed using SPM12 (Friston et al., 1994). In each subject, scans were spatially realigned and unwarped (using the first volume as a reference), normalised to MNI standard space using the deformation field from the unified segmentation/normalisation of the anatomical image, resampled to $2mm^3$ isotropic voxels, and smoothed with a 3D-Gaussian kernel of $8mm^3$ FWHM. Each voxel’s time series data was high-pass filtered at 1/128 Hz. Stimuli were modelled in an event-related fashion: after convolution with a canonical haemodynamic response function, which modelled the onset and duration of each sentence, separate regressors were entered into the general linear model (GLM) for each condition. Null events with and without babble noise were modelled as separate regressors. As well as the twelve experimental conditions (A, V, AVsync, AV-250, AV+250, null; each in the context of babble and no babble) we included six realignment parameters in the GLM for each participant to account for residual effects of motion. Once the model was estimated, we created *t*-contrasts for each condition to be passed to the second level. These contrasts indexed mean activation in each condition with the relevant null/baseline activation (babble or non-babble) subtracted. Resulting contrast images were entered into a second-level ANOVA for random effects analysis. This ANOVA modelled 24 conditions (twelve each for the younger and older age groups).

At the random effects level we implemented *t*-contrasts to test various effects. First, we tested for overall age differences in activation, in each direction, for all stimulus conditions separately. For example, greater activation in younger than older adults for auditory stimuli presented in the context of babble noise was assessed using $Y(Aud_{Babble}) > O(Aud_{Babble})$. We then tested for interactions between age and asynchrony in the context of babble noise using the contrasts $Y(AVasync_{Babble} > AVsync_{Babble}) > O(AVasync_{Babble} > AVsync_{Babble})$ and $O(AVasync_{Babble} > AVsync_{Babble}) > Y(AVasync_{Babble} > AVsync_{Babble})$. Results from these analyses are reported with a voxel threshold of $p < .001$, *FWE*-corrected at the cluster level at $p < .05$ with an extent threshold of > 100 voxels.

Finally, to locate mSTC, we tested the conjunction $(Aud_{NoBabble}) \cap (Vis_{NoBabble})$ with a conservative conjunction null (Friston, Penny, & Glaser, 2005). After applying stringent threshold criteria, specifically whole-brain voxel-level *FWE* correction at $p < .05$ and a cluster extent threshold of > 500 voxels, three clusters remained. Two that corresponded to STC in each hemisphere, and a third response-related cluster centred on left motor cortex; we manually selected the two STC clusters. We first carried out this analysis separately in each age group, to investigate age differences in the extent of mSTC. This was then repeated across all participants to define an ROI for masking the following analyses. Aud > Vis and Vis > Aud contrasts were used to test presence and location of sensory dominance along mSTC in each age group, in the context of silence and background noise. We then repeated the synchrony x age interaction contrasts, $\mathrm{Y}(\mathrm{Async}_{\mathrm{Babble}} > \mathrm{Sync}_{\mathrm{Babble}}) > \mathrm{O}(\mathrm{Async}_{\mathrm{Babble}} > \mathrm{Sync}_{\mathrm{Babble}})$ and $\mathrm{O}(\mathrm{Async}_{\mathrm{Babble}} > \mathrm{Sync}_{\mathrm{Babble}}) > \mathrm{Y}(\mathrm{Async}_{\mathrm{Babble}} > \mathrm{Sync}_{\mathrm{Babble}})$, to test for these effects specifically within mSTC. Results from these ROI analyses are reported with a voxel threshold of $p < .001$, *FWE*-corrected at the cluster level with small-volume correction and no minimum cluster size. To increase statistical power, we pooled over asynchrony levels in all fMRI analyses.

Resting-state fluctuation amplitude correction

Cortical blood flow and tissue perfusion has been shown to vary as a function of age (Riecker et al., 2003). This can make between-group comparisons of BOLD activation difficult, as any differences may simply reflect these physiological factors and not age-related changes in activation. To account for this problem, we applied a resting-state fluctuation amplitude (RSFA) correction to each participant's individual parameter estimate images (Kannurpatti & Biswal, 2008; Tsvetanov et al., 2015). Individual RSFA maps were calculated for each participant based on activity during the resting-state run. These data were first preprocessed in the same way as the task-related functional data. The following extra steps were then applied: wavelet despiking (Patel et al., 2014) and regression of realignment parameters (and their first derivative) to minimise the influence of movement on the signal, covarying out of mean white matter and cerebrospinal fluid signal (identified based upon SPM12 probability maps thresholded at 0.75), linear and quadratic detrending of the time series, and application of a 0.01 – 0.08 Hz bandpass filter. The standard deviation of the remaining signal was taken at each voxel to create the RSFA map. Values in each first-level parameter estimate image were then divided by the values in this map.

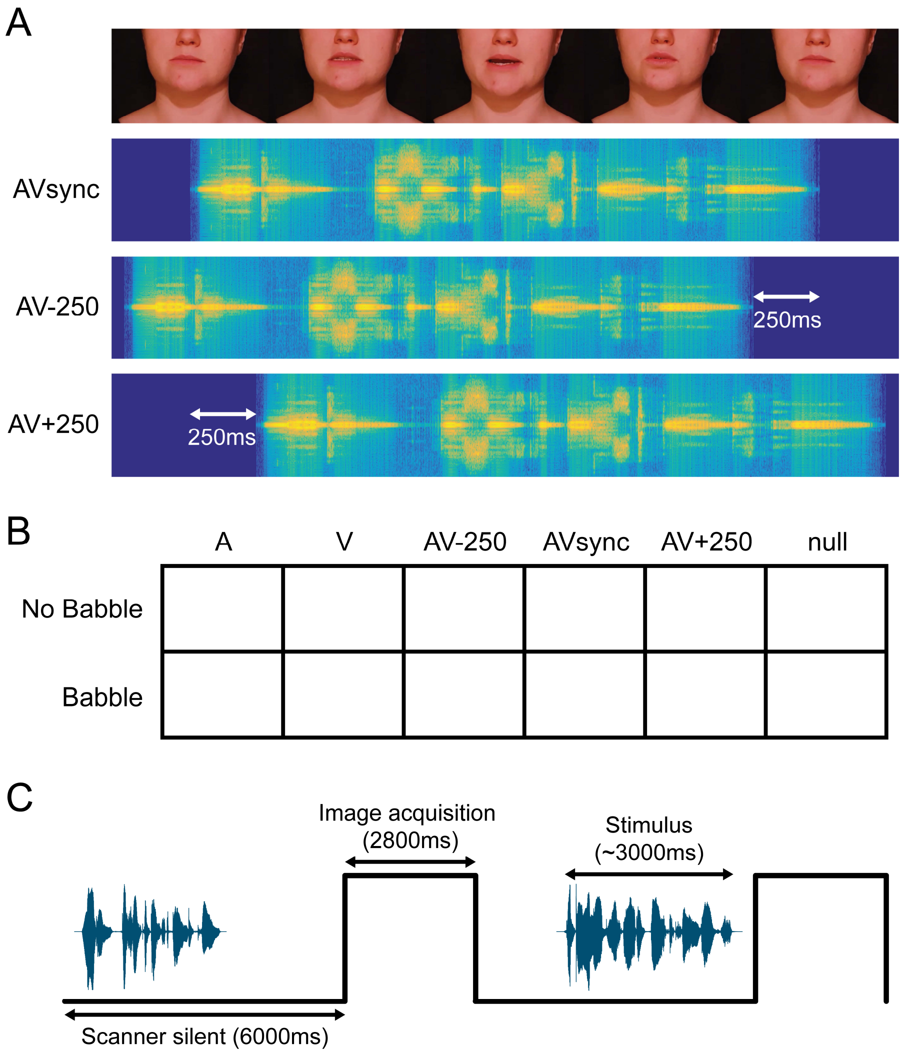

Figure 5.1. Experimental paradigm. (A and B) In the scanner (and the preceding dual-response behavioural session) participants were presented with spoken sentences under silent conditions or with constant 5-talker babble at 0dB SNR. Sentences could be auditory only (A), video only (V), synchronous auditory and video (AVsync), or asynchronous with auditory leading (AV-250) or trailing (AV+250) the video by 250ms. Including null trials presented inside the scanner, the paradigm therefore formed a 6 (stimulus type) x 2 (noise level) factorial design. Participants rated their comprehension via button press (behavioural and fMRI sessions) and repeated the sentence to the experimenter (behavioural session only). (C) A temporally-sparse imaging sequence was used to prevent MRI scanner noise from disrupting comprehension. Each volume acquisition was followed by a 6000ms silent period, during which a stimulus was presented with a randomly jittered onset. During babble runs, the babble track was played uninterrupted throughout.

Results

Audiometry

At the beginning of the screening session we tested participants' pure tone hearing thresholds. The summary of results in Figure 5.2 shows that older adults had mild impairments relative to the younger group at frequencies of 2000 Hz and below, and more severe hearing loss at 4000 and 8000 Hz. A mixed ANOVA of thresholds (averaged between ears in each participant) confirmed this with main effects of age, $F(1, 28) = 63.70$, $p < .001$, $\eta^2_p = .695$, and frequency, $F(2.864, 80.188) = 51.52$, $p < .001$, $\eta^2_p = .648$, as well as a strong interaction, $F(2.864, 80.188) = 29.36$, $p < .001$, $\eta^2_p = .512$.

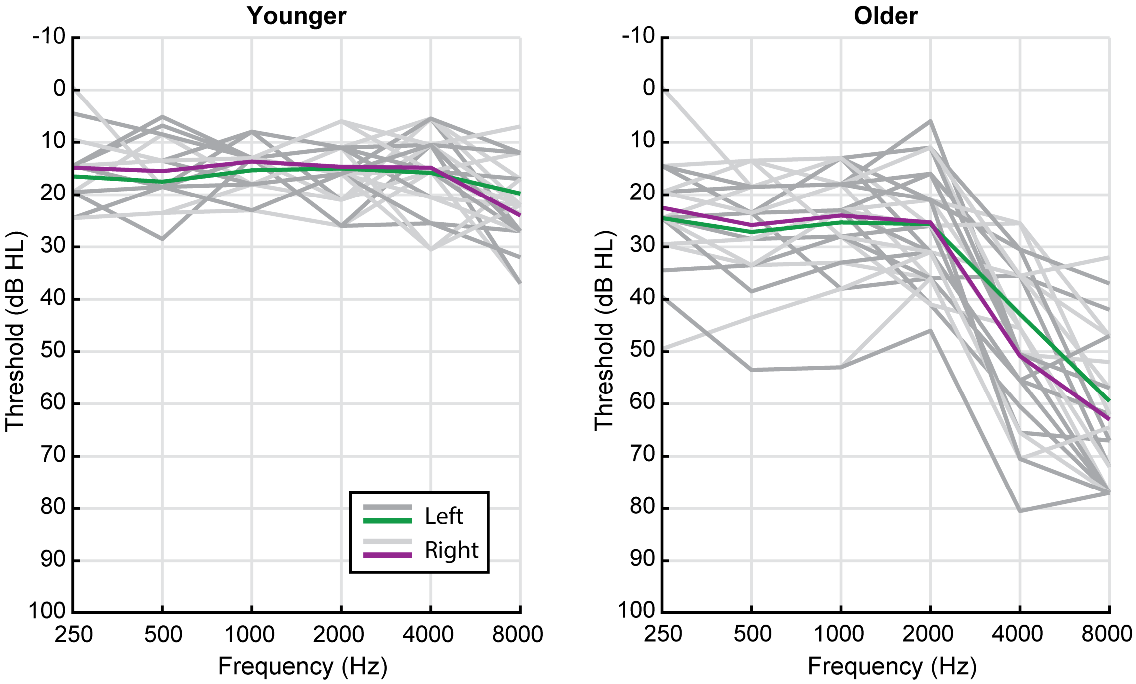

Figure 5.2. Audiograms for younger and older participants. Grey bars show individual thresholds; coloured bars show group means for each ear individually.

Synchrony judgement

To assess whether age or background noise affected participants' ability to explicitly judge the synchrony of spoken sentence stimuli, we presented audiovisual sentences synchronously and under four levels of asynchrony (-250, -150, 150, and 250 ms), with and without babble noise. For plotting, responses were quantified in terms of the probability of responding "same source" under each condition. For analysis, these were transformed into accuracy scores by taking the complementary probability of all scores except those in the synchronous conditions. As anticipated, there was a strong effect of synchrony on responses, $F(2.369, 61.598) = 133.753$, $p < .001$, $\eta^2_p = .837$. The presence of babble noise did not significantly affect responses, however, $F(1, 26) = 1.809$, $p = .190$, $\eta^2_p = .0.065$. Neither did age, though there was an obvious trend of older adults performing slightly better in most conditions, $F(1, 26) = 3.477$, $p = .074$, $\eta^2_p = .118$. Simple main effects show that older adults were significantly more accurate at the -150 ms asynchrony level, $p = .045$, but at no others. Age did not interact with noise, $F(1, 26) = 0.468$, $p = .500$, $\eta^2_p = .0.018$, or synchrony, $F(2.369, 61.598) = 2.197$, $p = .111$, $\eta^2_p = .0.078$. A small synchrony x noise interaction was apparent, $F(2.555, 66.423) = 3.996$, $p = .015$, $\eta^2_p = .133$, though this is difficult to interpret as it does not follow a clear shape. The three-way interaction was not significant, $F(2.555, 66.423) = 0.378$, $p = .737$, $\eta^2_p = .0.014$.

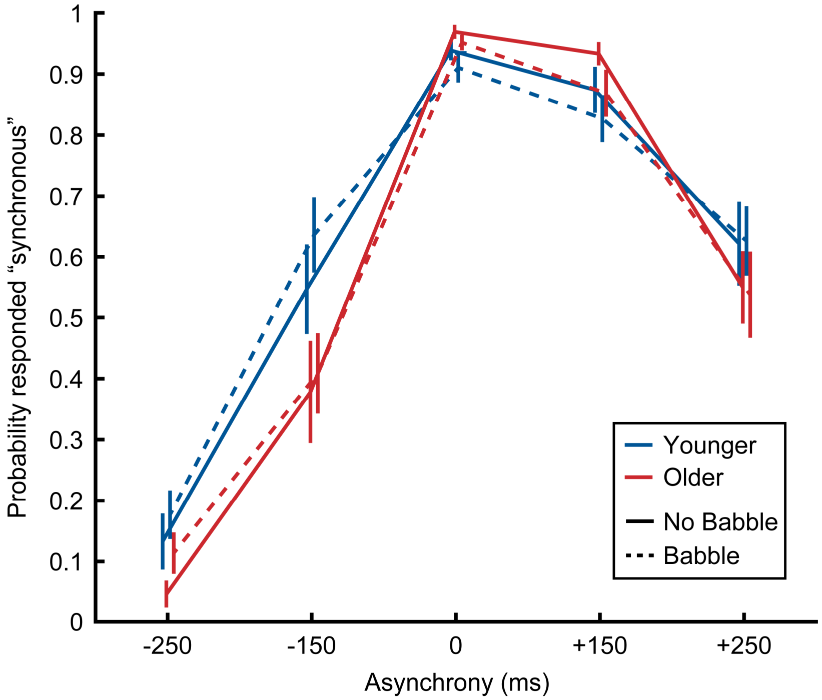

Figure 5.3. Results of the synchrony judgement task. Lines show mean probability of responding "same source" across participants. Negative asynchrony values indicate that the auditory preceded the visual. Error bars show ±1 *SEM*.

Dual-response paradigm

During the behavioural session, the younger and older participants also took part in a dual-response paradigm that assessed how well they could comprehend audiovisual sentences presented under conditions of asynchrony and noise. This task also served to confirm that the button-press response used inside the scanner correlated well with a more direct measure of comprehension. After each stimulus, participants first rated their comprehension with a button press, then repeated the sentence to the experimenter who scored it out of five based on target words. The results are summarised in Figure 5.4A.

Button response

In contrast to the synchrony judgement, older adults performed significantly more poorly overall, evidenced by a main effect of age, $F(1, 28) = 17.75$, $p < .001$, $\eta^2_p = .388$. The presence of background noise also significantly impaired responses, $F(1, 28) = 359.097$, $p < .001$, $\eta^2_p = .928$. There was also a large main effect of synchrony, $F(2.122, 59.411) = 1215.870$, $p < .001$, $\eta^2_p = .977$, with both groups performing best when stimuli were synchronous and less well under unimodal auditory and asynchronous conditions. It is clear from Figure 5.4, however, that this variability associated with age and asynchrony occurs only under the context of background noise, as both age groups perform at ceiling levels for all conditions (except visual alone) when no background noise is present. This is confirmed by noise x age, $F(1, 28) = 8.046$, $p = .008$, $\eta^2_p = .223$, noise x synchrony, $F(4, 112) = 102.770$, $p < .001$, $\eta^2_p = .786$, and the three-way noise x synchrony x age, $F(4, 112) = 12.959$, $p < .001$, $\eta^2_p = .316$, interactions. Synchrony also interacted with age, $F(2.122, 59.411) = 3.160$, $p = .047$, $\eta^2_p = .101$.

A small ANOVA focusing on the comparison between synchronous and (pooled) asynchronous stimuli in the noisy condition showed, via an age x synchrony interaction, that older adults had a stronger asynchrony effect, $F(1, 28) = 15.54$, $p < .001$, $\eta^2_p = .362$. A similar ANOVA comparing the auditory with audiovisual conditions confirmed the presence of significant multisensory enhancement, $F(1, 28) = 258.271$, $p < .001$, $\eta^2_p = .902$, which the interaction revealed was slightly stronger in the older age group, $F(1, 28) = 6.926$, $p = .013$, $\eta^2_p = .199$.

Sentence repetition

The pattern of results was very similar in the sentence repetition task. Main effects of age, $F(1, 28) = 16.14$, $p < .001$, $\eta^2_p = .366$, noise, $F(1, 28) = 247.472$, $p < .001$, $\eta^2_p = .898$, and synchrony, $F(2.762, 77.325) = 2091.962$, $p < .001$, $\eta^2_p = .987$, were all present. All interactions were also significant: noise x age, $F(1, 28) = 15.674$, $p < .001$, $\eta^2_p = .359$; synchrony x age $F(2.762, 77.325) = 5.154$, $p = .003$, $\eta^2_p = .155$; noise x synchrony, $F(3.103, 86.875) = 98.115$, $p < .011$, $\eta^2_p = .778$; and the three-way interaction, $F(3.103, 86.875) = 10.972$, $p < .001$, $\eta^2_p = .282$.

The smaller comparison between synchronous and pooled asynchronous conditions in the presence of noise again showed that older adults had a stronger asynchrony effect, $F(1, 28) = 12.65$, $p = .001$, $\eta^2_p = .311$. There was also a strong multisensory enhancement effect across groups, $F(1, 28) = 201.764$, $p < .001$, $\eta^2_p = .878$. The only way in which sentence repetition noticeably diverged from button-responses was that the small difference in multisensory enhancement between groups was not apparent here, $F(1, 28) = 1.586$, $p = .218$, $\eta^2_p = .054$.

Correlation between responses

The correlation between responses in the dual-response paradigm was excellent; the lowest score in either group was $r = .860$ (from a participant in the younger age group). There was no difference between the younger ($M = .932$, $SD = 0.038$) and older ($M = .935$, $SD =$

0.036) participants in the size of this correlation, $t(28) = 0.226$, $p = 0.823$, $d = 0.082$. This confirmed that the button-response task provided a good measure of sentence comprehension.

Behavioural results in scanner

Inside the scanner, participants performed the same button-press evaluation of sentence comprehension as in the dual-response paradigm. These results are summarised in Figure 5.4B. As our sparse imaging sequence allowed for the elimination of gradient noise during stimulus presentation we expected in-scanner behavioural results to closely resemble those in the screening session. This was the case, as confirmed by main effects of age, $F(1, 28) = 9.297$, $p = .005$, $\eta^2_p = .249$, noise $F(1, 28) = 200.741$, $p < .001$, $\eta^2_p = .878$, and synchrony, $F(2.695, 75.447) = 1460.391$, $p < .001$, $\eta^2_p = .981$, as well as the noise x age, $F(1, 28) = 8.787$, $p = .006$, $\eta^2_p = .239$, noise x synchrony, $F(2.306, 64.564) = 75.375$, $p < .001$, $\eta^2_p = .729$, and three-way, $F(2.306, 64.564) = 5.901$, $p = .003$, $\eta^2_p = .174$ interactions.

In this instance synchrony and age did not significantly interact within the main ANOVA, $F(2.695, 75.447) = 1.713$, $p = .176$, $\eta^2_p = .058$. However, the smaller analysis that pooled over asynchrony levels and focused only on bimodal trials in the context of babble showed that older adults still had a stronger asynchrony effect inside the scanner, $F(1, 28) = 8.367$, $p = .007$, $\eta^2_p = .230$. See the bottom right panel of Figure 5.7B for a plot of this interaction. The second sub-analysis again confirmed the presence of a strong multisensory enhancement effect, $F(1, 28) = 78.592$, $p < .001$, $\eta^2_p = .737$. In agreement with the sentence repetition data, this did not interact with age, $F(1, 28) = 0.779$, $p = .385$, $\eta^2_p = .027$.

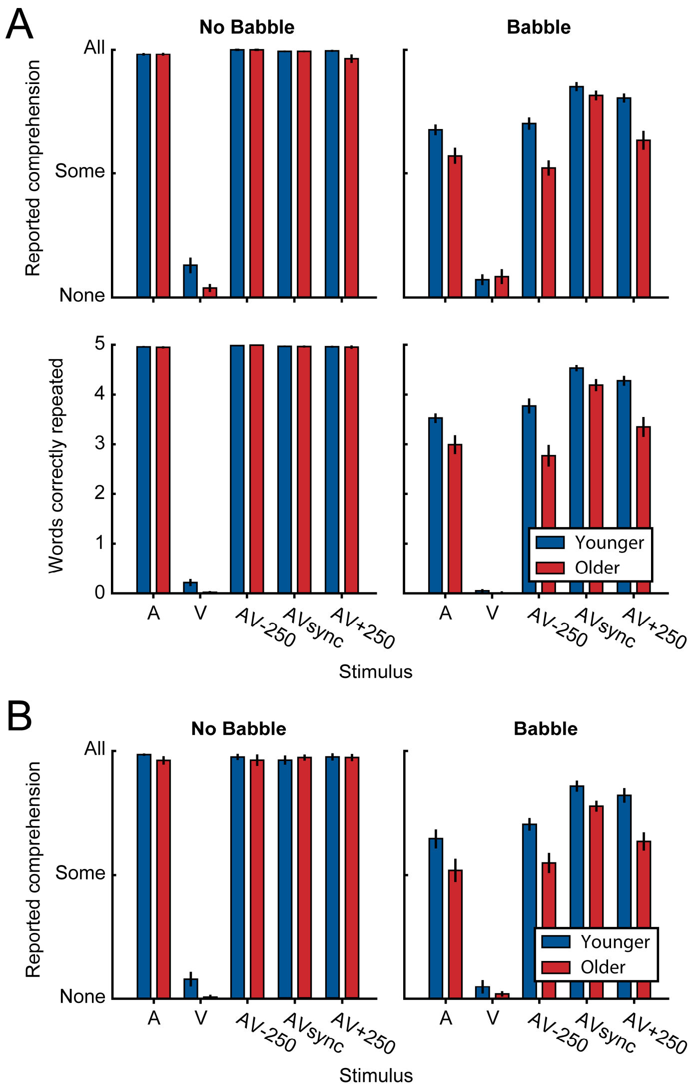

Figure 5.4. Sentence comprehension results. (A) In a dual-response paradigm during the behavioural testing session, participants first reported comprehension of each sentence via button press (top row), then repeated the sentence back to the experimenter who assigned a score out of five based on target words (second row). (B) During scanning participants responded only with a button press. Negative asynchrony values indicate that the auditory preceded the visual. Bars show group means ±1 *SEM*.

fMRI results

Overall age differences

We first compared overall task-related activation in younger and older adults. Results are shown in Figure 5.5. Under quiet conditions, in which the groups were behaviourally matched, younger adults showed significantly higher bilateral activation centred on superior temporal gyrus and extending to posterior insula in response to all stimuli with an auditory component ($A_{NoBabble}$, $AVsync_{NoBabble}$, and $AVasync_{NoBabble}$). The pattern was similar across the three conditions. Older adults only showed significantly greater activation in response to unimodal stimuli, particularly for $V_{NoBabble}$ for which there were significant differences bilaterally in visual areas including middle occipital gyrus and intraparietal sulcus. The older group also had greater superior frontal activations, significant only in unimodal conditions, that were presumably task-related.

In the context of background noise, where older adults indicated overall lower comprehension and a greater disruptive effect of asynchrony, the younger age group only showed significantly greater activations in response to the more challenging auditory conditions (A_{Babble} and $AVasync_{Babble}$). These were again centred on superior temporal gyrus and insula, but in this case left-lateralised and more extensive, extending in the $AVasync_{Babble}$ condition to supramarginal gyrus. In contrast to the quiet conditions, when babble noise was present older adults showed significantly greater visual activations in response to bimodal stimuli, particularly in the AV_{sync} condition where a large bilateral network of visual regions extending from middle occipital gyrus to anterior intraparietal sulcus was apparent.

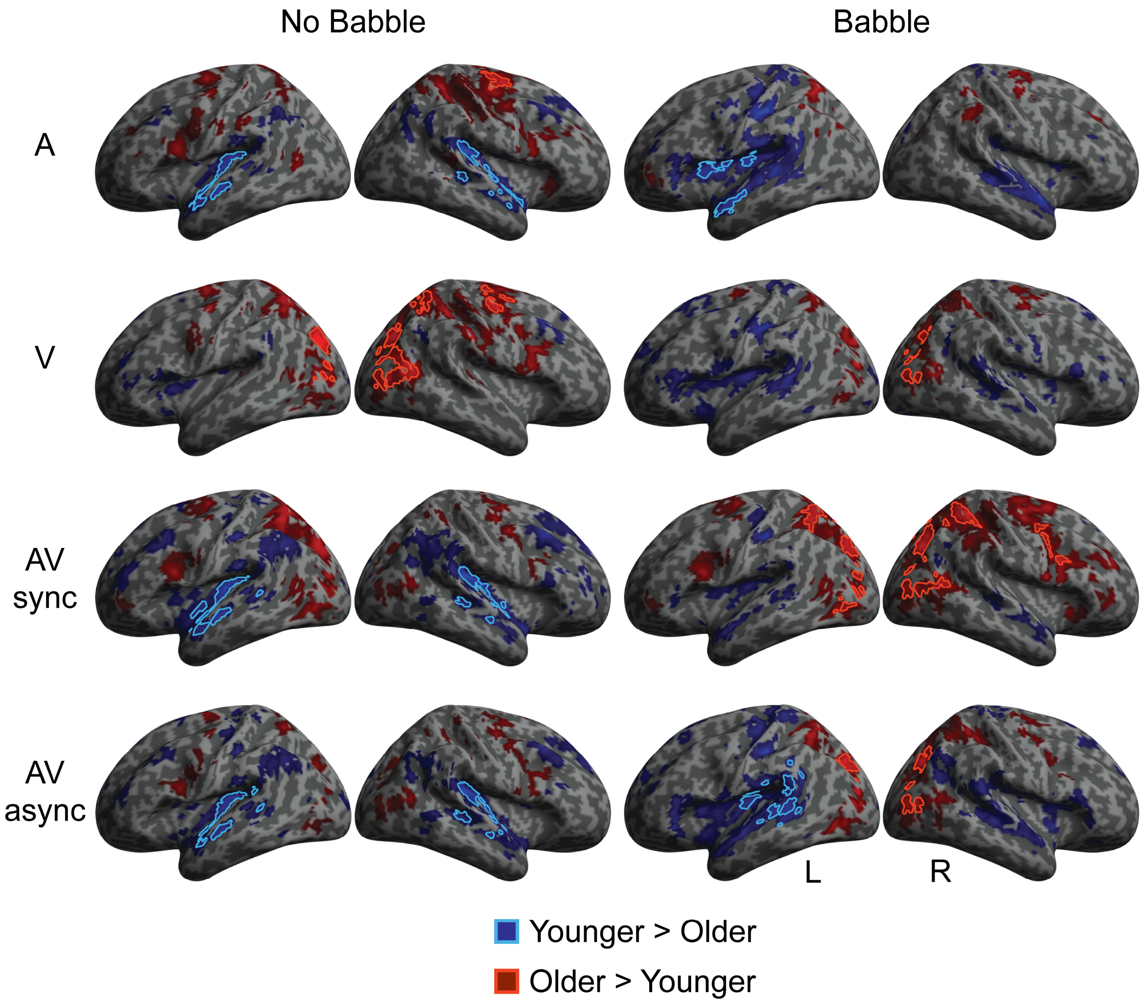

Figure 5.5. Age-related activation differences under each stimulus condition are shown rendered on an inflated template brain. Activations are thresholded at $p < .05$ uncorrected, extent threshold > 100 voxels, for illustration. Clusters significant at $p < .05$ with *FWE* correction, activation threshold $p < .001$, extent threshold > 100 voxels, are highlighted in brighter colours. Note that large areas of activation may appear as groups of smaller clusters on the inflated brain.

Interaction between age and synchrony

As the behavioural data showed that older adults were more affected by stimulus incongruence in the context of babble, we next investigated the interaction between age and stimulus synchrony. The contrast Y(Async$_{\text{Babble}}$ > Sync$_{\text{Babble}}$) > O(Async$_{\text{Babble}}$ > Sync$_{\text{Babble}}$) revealed an array of significant age differences in superior frontal areas centred on supplementary motor area (peak at [-8 2 66], *z*-score = 5.04), anterior cingulate cortex (peak at [6 20 34], *z*-score = 4.71), right postcentral gyrus (peak at [30 -30 66, *z*-score = 4.69]), right inferior frontal gyrus (peak at [26 22 -8], *z*-score = 4.53), and bilateral anterior insular cortex (left peak at [-40 12 -10], *z*-score = 4.09; right peak at [34 20 8], *z*-score = 4.14). See Figure 5.6 for an overview. The opposite contrast O(Async$_{\text{Babble}}$ > Sync$_{\text{Babble}}$) > Y(Async$_{\text{Babble}}$ > Sync$_{\text{Babble}}$) produced no significant results.

Multisensory superior temporal cortex

Finally, we were interested in the extent and activity of multisensory superior temporal cortex (mSTC) in younger and older adults. First, we defined the area (superior temporal regions that are significantly active for the conjunction $[\text{Aud}_{\text{NoBabble}} > \text{Baseline}_{\text{NoBabble}}] \cap [\text{Vis}_{\text{NoBabble}} > \text{Baseline}_{\text{NoBabble}}]$) separately in each age group for comparison. Results are shown in Figure 5.7A. The regions mostly overlap between the age groups, though their extent is slightly greater in younger participants, particularly on the left where mSTC extends further anteriorly towards temporal pole. We then defined the region using the same contrast across all participants to use as an ROI mask in the following analyses.

To assess the distribution of auditory and visual dominance along mSTC we used simple Aud > Vis and Vis > Aud contrasts separately for each age group and each noise condition. Results can be seen in Figure 5.7B. Again, the groups' activations mostly overlapped but the area was reduced in older adults, especially in the presence of babble noise. In agreement with previous research, in all cases the more anterior parts of mSTC showed greater auditory activation. In a quiet context, the posterior regions showed no significant activation for either contrast, implying that they were activated to a similar degree for auditory and visual stimuli. In the presence of babble noise the areas of auditory dominance extended less posteriorly, while the V > A contrast revealed significant clusters in right posterior mSTC in both age groups.

Finally, to explore whether mSTC encoded stimulus asynchrony differently between age groups, we tested the age x synchrony interactions described above specifically within this region. The $\text{Y}(\text{Async}_{\text{Babble}} > \text{Sync}_{\text{Babble}}) > \text{O}(\text{Async}_{\text{Babble}} > \text{Sync}_{\text{Babble}})$ contrast revealed a cluster in right posterior mSTC with increased activation for asynchronous relative to synchronous stimuli in the younger age group, and vice-versa in the older group, only in the context of babble noise. See Figure 5.7C for cluster location and a comparison between parameter estimates and behavioural responses.

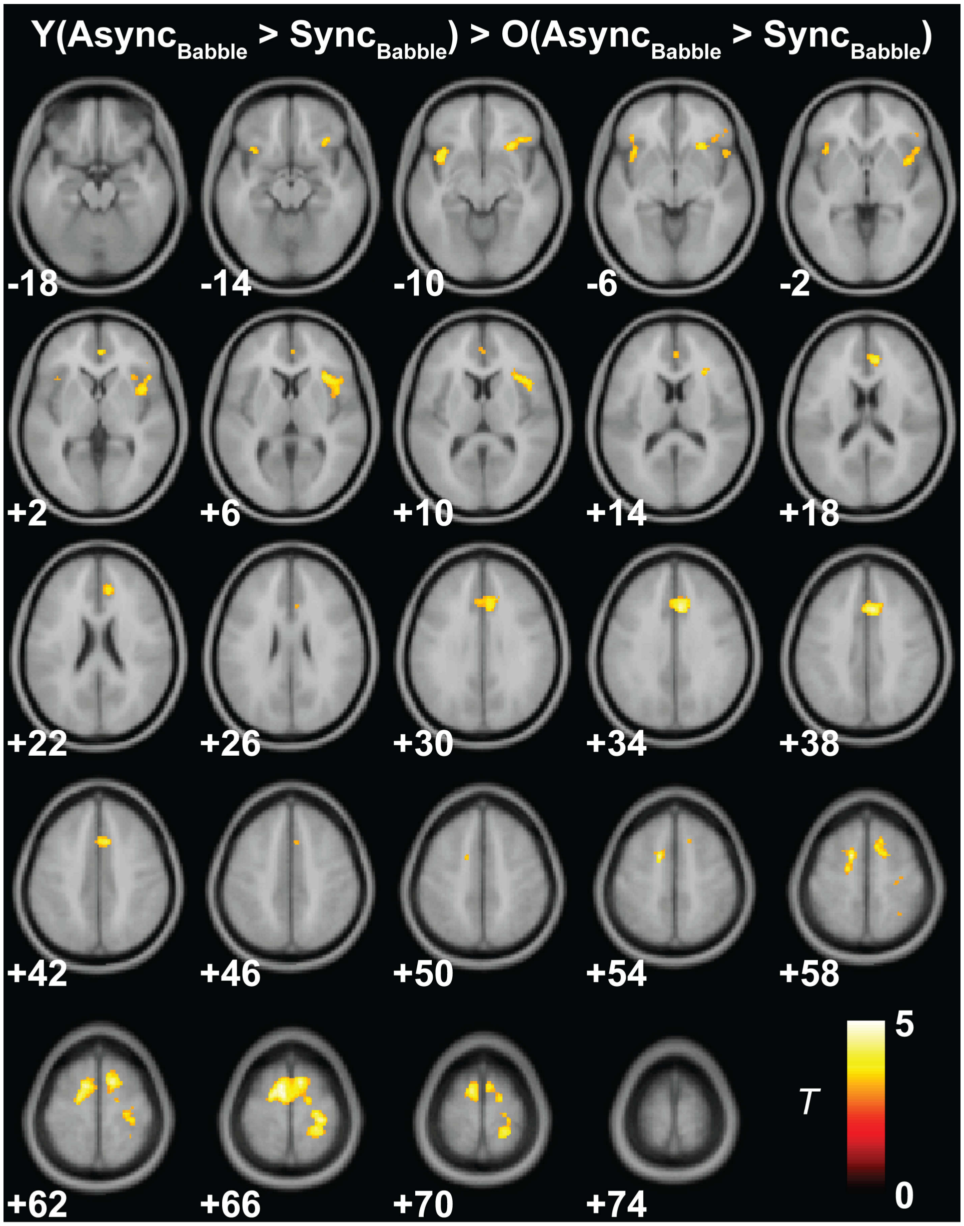

Figure 5.6. Greater activation for younger adults in response to asynchronous versus synchronous stimuli, in the context of background babble noise. Results are shown on an average template brain, thresholded at $p < .001$, extent threshold > 100 voxels.

Table 5.1.

fMRI results: Group-level GLM

Region	Coordinates			z-score peak	Number of voxels	p value*
No Babble A (Y-O)						
L. superior temporal gyrus	-30	-34	18	6.26	818	0.000
L. superior temporal gyrus	-48	-12	-8	5.69		
L. transverse temporal gyrus	-34	-26	4	5.02		
R. superior temporal gyrus	42	-24	0	6.14	671	0.000
R. superior temporal gyrus	40	-34	16	6.02		
R. superior temporal gyrus	54	-8	-2	4.09		
No Babble A (O-Y)						
R. precentral sulcus	24	-12	54	4.71	286	0.006
R. precentral sulcus	26	-16	72	3.32		
R. supplementary motor area	4	6	62	4.13	261	0.009
R. supplementary motor area	4	-4	52	4.08		
No Babble V (O-Y)						
R. middle occipital gyrus	36	-80	12	5.57	1119	0.000
R. intraparietal sulcus	28	-68	30	5.15		
R. angular gyrus	50	-66	14	4.84		
L. intra-occipital sulcus	-24	-76	26	5.45	488	0.000
L. lateral occipital sulcus	-34	-82	12	4.74		
L. middle occipital gyrus	-44	-78	12	3.42		
R. fusiform gyrus	28	-46	-14	4.60	187	0.034
R. fusiform gyrus	34	-38	-18	3.87		
R. precentral sulcus	42	-6	54	4.14	293	0.005
R. precentral sulcus	28	-10	68	3.80		
R. precentral gyrus	44	-14	58	3.74		
No Babble AVsync (Y-O)						
L. superior temporal gyrus	-30	-32	18	6.12	868	0.000
L. superior temporal gyrus	-34	-26	4	5.38		
L. superior temporal sulcus	-48	-18	-8	4.83		
R. superior temporal gyrus	44	-22	-2	5.80	721	0.000
R. superior temporal gyrus	40	-34	16	5.62		
R. superior temporal gyrus	36	-26	8	4.31		

No Babble AVasync (Y-O)						
L. superior temporal gyrus	-32	-34	20	6.39	864	0.000
L. superior temporal gyrus	-34	-26	4	5.00		
L. superior temporal sulcus	-50	-18	-6	4.68		
R. superior temporal gyrus	40	-34	16	5.91	763	0.000
R. superior temporal gyrus	44	-24	0	5.88		
R. superior temporal sulcus	54	-8	-4	4.10		
Babble A (Y-O)						
L. superior temporal sulcus	-50	-6	-14	4.55	209	0.022
L. superior temporal gyrus	-50	8	-18	4.13		
L. anterior insula	-34	8	8	4.39	361	0.002
L. posterior insula	-40	-14	12	4.08		
L. posterior insula	-38	-4	12	3.63		
Babble V (O-Y)						
R. middle occipital gyrus	36	-82	16	5.38	216	0.019
Babble AVsync (O-Y)						
R. intraparietal sulcus	28	-70	32	5.70	1175	0.000
R. middle occipital gyrus	36	-78	10	5.66		
R. angular gyrus	50	-66	12	4.07		
R. superior parietal lobule	34	-50	58	4.65	485	0.000
R. intraparietal sulcus	26	-58	62	3.91		
R. postcentral gyrus	34	-42	50	3.44		
R. inferior frontal sulcus	38	6	28	4.43	304	0.004
R. precentral sulcus	44	-2	46	3.75		
R. precentral sulcus	46	-4	54	3.61		
L. intraparietal sulcus	-32	-56	56	4.39	433	0.001
L. intraparietal sulcus	-22	-56	48	3.67		
L. precuneus	-8	-64	64	3.48		
L. lateral occipital sulcus	-34	-84	12	4.33	183	0.037
L. lateral occipital sulcus	-44	-80	2	3.66		
L. lateral occipital sulcus	-44	-72	-2	3.57		
L. intraparietal sulcus	-24	-78	36	4.19	205	0.024
Babble AVasync (Y-O)						
L. supramarginal gyrus	-54	-44	26	4.75	561	0.000

L. superior temporal sulcus	-42	-42	6	4.22		
L. superior temporal sulcus	-52	-40	8	4.18		
Babble AVasync (O-Y)						
R. middle occipital gyrus	36	-78	10	5.73	663	0.000
R. intraparietal sulcus	28	-68	30	5.24		
R. intraparietal sulcus	30	-80	26	3.65		
L. intraparietal sulcus	-26	-76	28	4.20	233	0.014
Y(AVasyncBabble > AVsyncBabble) > O(AVasyncBabble > AVsyncBabble)						
L. superior frontal gyrus	-8	2	66	5.04	1070	0.000
R. superior frontal gyrus	12	6	64	4.80		
L. superior frontal gyrus	-14	4	58	4.77		
R. anterior cingulate cortex	6	20	34	4.71	439	0.000
R. central sulcus	30	-30	66	4.69	334	0.002
R. postcentral gyrus	24	-40	68	4.48		
R. precentral gyrus	24	-24	66	4.47		
R. anterior insula	26	22	-8	4.53	177	0.041
R. inferior orbital gyrus	34	28	-14	3.71		
R. inferior frontal gyrus	46	32	0	3.32		
R. anterior cingulate cortex	8	36	20	4.33	166	0.051
R. anterior cingulate cortex	2	44	2	3.79		
R. anterior insula	34	20	8	4.14	443	0.000
R. anterior insula	42	16	8	4.05		
R. anterior insula	42	8	2	3.95		
L. anterior insula	-40	12	-10	4.09	194	0.030
L. anterior insula	-34	18	-12	3.89		
L. inferior frontal gyrus	-38	30	-6	3.52		

**p*-values corrected for familywise errors at the cluster level, activation threshold = $p < .001$ uncorrected, extent threshold = 100 voxels

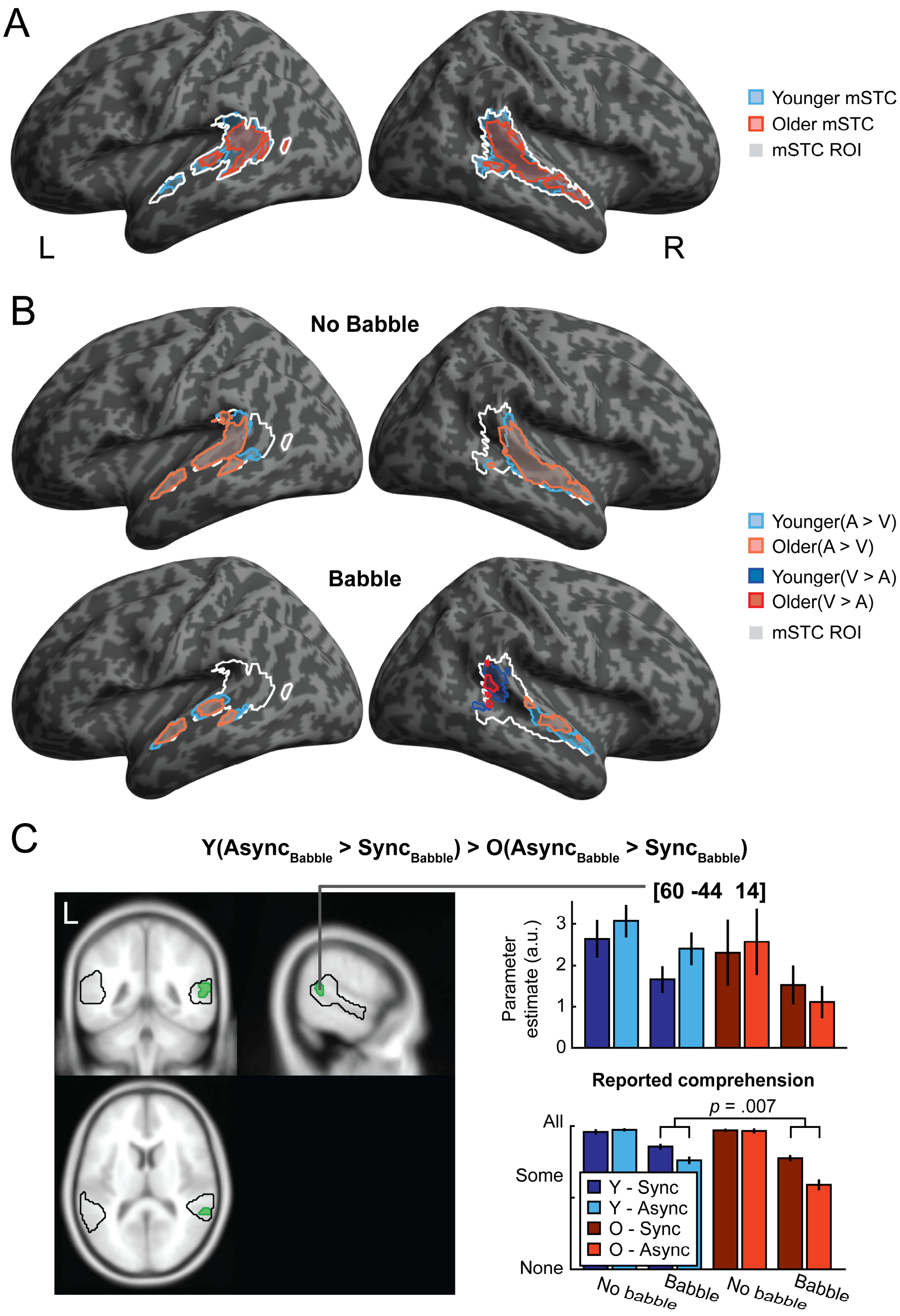

Figure 5.7. Multisensory superior temporal cortex (mSTC) analyses. (A) mSTC was defined as the conjunction $(\text{Aud}_{\text{NoBabble}} > \text{Baseline}_{\text{NoBabble}}) \cap (\text{Vis}_{\text{NoBabble}} > \text{Baseline}_{\text{NoBabble}})$. This was computed separately for younger and older adults for comparison, and on all participants together to use as the mSTC mask in analyses below. (B) Age differences in auditory and

visual dominance. (C) Age x asynchrony interaction in the context of babble. Results in B and C show clusters thresholded at $p < .05$ with small-volume FWE correction.

Discussion

Our rich behavioural data revealed both similarities and differences between younger and older participants. As anticipated, we observed that both age groups were less able to comprehend auditory spoken sentences when these were presented with masking background noise, and that this effect was stronger in older adults. We also found that the addition of a synchronous visual stimulus benefitted participants from both groups. In most cases this enhancement effect was of a similar size between age groups, though there was an indication of slightly greater improvement in the older group that only reached significance in the button responses of the dual-response paradigm. Audiovisual asynchrony led to significantly reduced comprehension compared to synchronous stimuli, though only in the presence of background noise; participants performed near perfectly under all noise-free conditions that contained an auditory signal. Younger adults showed stronger comprehension for asynchronous audiovisual relative to unimodal auditory stimuli, while in older adults these were mostly matched.

Intriguingly, the synchrony judgement data show a very different pattern. Neither age, nor the presence of background noise, significantly affected participants' ability to accurately report asynchrony in an audiovisual signal. In fact, older adults performed (non-significantly) better at this task under most conditions. This implies that the impairments in comprehension introduced by asynchrony in the presence of a masking background noise, and particularly the interaction with age, do not result from an inability to detect the asynchrony. This result echoes the findings of Bedard and Barnett-Cowan (2015), who showed that older adults were both less accurate at temporal order judgements and more susceptible to the stream-bounce illusion, but matched the performance of younger adults on explicit synchrony judgement tasks. The authors hypothesise that the neural mechanisms underlying the first two tasks are impaired with ageing, while those underlying explicit synchrony judgement (and which may sub-serve the others) are not. A similar explanation is possible in our case: detection of the timing difference is necessary, but not sufficient, to inhibit the interfering effects of an asynchronous visual signal. Higher-level mechanisms are perhaps also required to perform the necessary inhibitory/compensatory processes, and it is these that are selectively impaired in older adults.

This suggestion is supported by the interactions observed between synchrony and age in our fMRI data. An entire network of regions was more active for younger than older adults in response to asynchronous versus synchronous stimuli in the context of background noise. These regions, particularly anterior cingulate cortex and inferior frontal areas, correspond closely with the well-established cognitive control network that activates in response to incongruent stimuli (Fan, McCandliss, Fossella, Flombaum, & Posner, 2005; MacDonald, Cohen, Stenger, & Carter, 2000; Roberts & Hall, 2008). Age-related changes to this network have been observed previously (Mohtasib et al., 2012; Schulte et al., 2011; Sharp, Scott, Mehta, & Wise, 2006). In our case, the implication is that younger participants recruit this network to detect, evaluate, and perhaps compensate for asynchrony in the stimuli, facilitating improved behavioural responses.

Despite the dissociation between synchrony detection and comprehension performance, both measures showed asymmetrical effects of asynchrony. That is, participants were more likely to detect asynchrony, and less likely to comprehend the stimuli, when the auditory preceded the visual. This corresponds well with previous asynchrony detection (Conrey and Pisoni, 2006; Maier, Di Luca, & Noppeney, 2011) and speech comprehension (Grant, van Wassenhove, & Poeppel, 2004) research. To our knowledge this is the first study to show a direct impact of asynchrony on sentence comprehension in older adults (but see Başkent & Bazo, 2011, for a correlational study in hearing-impaired participants).

Neural data also revealed more general age differences. The younger adults' greater superior temporal activations in response to auditory speech is in agreement with the existing literature. Peelle, Troiani, Grossman, & Wingfield (2011) found a similar effect for older adults with mild to moderate presbycusis perceiving sentences in a quiet context, while Wong et al. (2009) and Hwang, Li, Wu, Chen, and Liu (2007) found the same for target speech stimuli embedded in masking noise. Our inclusion of visual stimuli that aid comprehension allows for a more nuanced understanding of this effect, however. We found that, in the absence of masking background noise, the greater auditory activation in younger adults was mostly unaffected by the presence or synchronicity of a visual stimulus. Conversely, when the auditory stimuli were embedded in noise, the effect was modulated by the presence and congruence of visual stimuli. Based on these findings, and conclusions of the previous work discussed above, it is probable that age differences in superior temporal responses to speech stimuli result from a combination of simple reduced cochlear input and other higher-level processes.

Increased visual activity in the older age group, present across multiple conditions but most apparent for synchronous stimuli presented in the context of background noise, suggests that older adults are more reliant on the visual component of an audiovisual speech signal than their younger counterparts. This also fits well with previous research. Electrophysiological data from Winneke & Phillips (2011) show that older adults exhibit earlier and more pronounced visual facilitation of speech comprehension, suggesting that they are more accomplished at extracting useful features from visual speech information. Neither we nor Winneke & Phillips observed behavioural differences between age groups in the degree of multisensory enhancement, but this is not necessarily expected: when audiovisual speech signals are perfectly congruent, much of the information is redundant across modalities (Park, Ince, Schyns, Thut, & Gross, 2018), meaning that a change in relative weighting of the signals need not result in a corresponding change in the final percept. Such differences would be more likely in the presence of incongruent stimuli, where an increase in the relative weighting of the visual cues may lead participants to respond in a visually-biased way. This has indeed been demonstrated previously in older adults' responses to McGurk stimuli (Cienkowski & Carney, 2002; Sekiyama, Soshi, & Sakamoto, 2014).

Finally, we compared the position and response profile of multisensory superior temporal cortex (mSTC) between younger and older adults. The extent of mSTC, defined as the regions of STC bilaterally that respond significantly to both unimodal auditory and unimodal visual speech stimuli, overlapped considerably between age groups. It covered a wider area, especially extending more anteriorly, than has been shown previously (Stevenson et al., 2010; Stevenson et al., 2011). This is probably due to a combination of the sparse sampling sequence, long, naturalistic sentence stimuli, and focused comprehension task making our paradigm more sensitive to these activations. We also found the anterior/posterior, auditory/visual divide that has been described previously (e.g. Ozker et al., 2017) in both age groups, though again the extent of significant activations was smaller in the older group.

In summary, this study revealed converging behavioural and neural age differences in response to synchronous and asynchronous audiovisual speech stimuli presented in quiet and noisy environments. Older adults were more affected by background noise and stimulus asynchrony, and showed a stronger visual bias in neural responses, but both age groups benefitted similarly from the presence of synchronous visual information. Further, we showed that mSTC is largely unchanged in older adults, and that detection of synchrony and audiovisual speech comprehension are differentially affected by healthy ageing.

CHAPTER 6: GENERAL DISCUSSION

In this thesis I have engaged with the problem of multisensory integration in the healthy ageing brain. A review of the current literature (Chapter 1) revealed a striking lack of consensus regarding the impact of healthy ageing on multisensory integration. Some behavioural studies, particularly those that measure multisensory enhancement of reaction times, report that multisensory information is more beneficial to older than younger adults. Others, especially those that require the suppression of an incongruent stimulus, find age-related impairments. In each case there are papers using the same or similar tasks that report null or opposite effects, alongside a great number of other approaches that show little to no difference between the age groups. Even when an effect is robust and highly replicable, such as older adults' greater susceptibility to the double-flash illusion, the underlying explanation remains elusive. Neuroimaging research is more limited but similarly mixed in its findings.

These findings made two points clear: First, there is a need for more mechanistic explanations of the age differences that are observed. This area of research is relatively young, and the nature of these differences are still being established, but it is important that we begin attempting to explain as well as describe them. More neuroimaging work, alongside the application of established computational models, can help with this aim. Second, any work addressing this problem must consider the specific demands of the behavioural tasks used and accept the possibility that the direction and magnitude of age differences may vary with stimuli and task demands. As I note briefly in Chapter 1, multisensory integration is not a single monolithic process, but a term used to describe a wide array of functions and behaviours that happen to involve the processing of information from multiple sensory modalities. Embracing this truism may allow us to approach a richer understanding of multisensory perception in older adults.

Overview

The empirical research described in Chapters 2 – 5 was conducted with these ideas in mind. My colleagues and I combined behavioural testing, computational modelling, and functional brain imaging with the aim of improving our understanding of the mechanisms underlying age-related changes in two different forms of audiovisual integration. I begin this final chapter with a brief account of each study and its implications.

Chapter 3: Older Adults Sacrifice Response Speed to Preserve Multisensory Integration

This study addressed multiple questions. First, and most basically, we hoped to establish how older adults respond to the ventriloquist illusion. The effects of ageing on perception of other multisensory illusions, including the sound-induced flash illusion, stream/bounce illusion, and McGurk effect, have received considerable attention, so it is surprising that (to my knowledge) nobody has previously tested older adults with ventriloquist stimuli. As this illusion is frequently investigated in the younger population and measures a process that is ubiquitous in daily living (multisensory spatial localisation and orienting), information about any age differences are interesting on their own. This illusion is also useful because it can be considered conceptually opposite to the sound-induced flash illusion (one illustrates visual capture of spatial information, the other auditory capture of temporal information) and relies on separate mechanisms (detection versus localisation), despite using similar transient low-level stimuli. Differences in age-related effects between these tasks may thus reveal fundamental information about the specific features of multisensory stimuli that lead to increased integration in older adults. We characterised responses using both explicit and implicit measures of integration (common-source judgement versus auditory localisation response) because previous research has indicated that age may affect these differently (e.g. Bedard & Barnett-Cowan, 2015).

Our findings were unequivocal: older adults were no more susceptible to the ventriloquist illusion than their younger counterparts. Across multiple levels of visual reliability and spatial separation, the age groups consistently gave the same common-source and auditory localisation responses. The only apparent difference between groups was in reaction times to the common-source judgement task, which indicated that older adults took disproportionately longer to respond to more ambiguous stimuli (i.e. where visual reliability was low or spatial separation small, leading to greater uncertainty about the probability of the stimuli sharing a source).

The use of the ventriloquist task to measure multisensory integration was further motivated by our second question: does the best-performing computational model of multisensory integration in young adults also predict the behaviour of healthy older adults, and are the fitted parameters similar between groups? The BCI model has been used to model responses to many types of multisensory stimuli, but a large proportion of the work in developing and refining the model has involved the ventriloquist effect (Koerding et al., 2007; Rohe & Noppeney, 2015a; Shams & Beierholm, 2010; Wozny, Beierholm, & Shams, 2010), making this the natural choice for a first attempt to fit the model in a new population. Our aim was to apply a standard form of the BCI model to younger and older participants' localisation responses and compare the fitted parameters.

In agreement with the model-free analyses, the BCI model fitted equally well to responses from both younger and older adults, with parameters that did not differ between age groups. Considered individually, these parameters showed that age had no significant effect on participants' prior tendency to bind signals, nor on the strength of their spatial prior, nor the reliability of their sensory representations. This final point was further supported by participants' unimodal localisation responses, which revealed very few group differences in auditory and visual localisation performance.

The BCI modelling and these unimodal localisation tests, alongside the multiple other tasks used, were also designed to address a third question: What mechanisms underlie any age differences found? As already noted, older and younger adults did not differ in their final responses to the self-speeded ventriloquist paradigm, so there was nothing to explain here. However, age did interact with reaction times on the common-source judgement task, which might suggest differences in the underlying cognitive processes. These results should be interpreted with caution as participants had not been instructed to respond quickly, but they were supported by the outcome of a simplified, speeded ventriloquist task in which participants also took part. In this case, as well as being slower overall (an effect that was present to some degree in every task described in this thesis), older adults were disproportionately slower at responding to the most challenging condition (locate auditory in the presence of a spatially incongruent visual distractor). To ensure that this was not simply a result of older adults using different speed/accuracy criteria, a well-established effect (Hertzog, Vernon, & Rypma, 1993; Rabbitt, 1979; Salthouse, 1979; Smith & Brewer, 1995), we fitted a Bayesian compatibility bias (evidence accumulation) model to the responses. This model revealed that older adults do indeed set a higher response threshold (i.e. prefer to be more certain before giving a response), but crucially also take longer than younger participants to reach a given level of certainty, suggesting that the evidence accumulation process becomes noisier with age.

In answer to this third question, we can therefore conclude that both age groups respond similarly to ventriloquist-type stimuli given unlimited time, but the decision process is slower and noisier in older adults. As these tasks used transient stimuli, the compatibility bias model did not measure evidence accumulation in the standard sense, as there was no evidence left to accumulate after 50 ms. Instead, it indexed some internal process that attempted to resolve noisy multisensory signals into a veridical representation of the stimuli, and which appears to change with age.

Chapter 4: Cortical Hierarchies for Multisensory Integration Are Preserved in Healthy Older Adults

It was this process that the study described in Chapter 4 aimed to explore. Previous research (Rohe and Noppeney, 2015a, 2015b) has demonstrated (in healthy young adults) that audiovisual integration of spatial information takes place across a cortical hierarchy, each region of which encodes stimuli according to a specific stage of the BCI model. However, investigations into the effects of ageing on neural activity have demonstrated increased cortical activation in older adults in response to a wide variety of tasks (Cabeza, 2002; Davis et al., 2008; Park and Reuter-Lorenz, 2009; Grady, 2012). This is often explained in terms of compensatory activation, though recent work has questioned this interpretation (Morcom & Johnson, 2015; Morcom et al., 2018). Our aim with this study was thus to ascertain whether the cortical hierarchy for audiovisual integration degrades with age, and to search for compensatory neural activity that may correspond to the behavioural differences seen in Chapter 3.

We again employed a ventriloquist paradigm for this study. Despite the task differing from that used in Chapter 3 in a number of important ways, including the presence of background noise and the use of simulated sound location cues and button responses, we replicated the behavioural results: younger and older adults performed similarly at unimodal and bimodal spatial localisation, and did not differ significantly in susceptibility to the ventriloquist effect. Neural responses during the ventriloquist task were analysed in three ways. First, we ran a standard mass-univariate analysis to identify age differences in task-related activation levels. Next, we applied multivariate Bayesian decoding analyses to quantify the type and quality of stimulus information represented in the regions that were activated more by the older adults, and compared this with regions activated by both groups. Finally, we used support vector regression machines to assess the type of multisensory spatial information held in pre-specified regions of the cortical hierarchy and compared this between groups.

We found that older adults did indeed activate an array of cortical regions to a greater degree than the younger age group while performing the task. However, the multivariate Bayesian analysis revealed no age differences in the stimulus information encoded in these areas, so the increased activity could not be interpreted as reflecting compensatory recruitment of extra regions. Furthermore, representations of the stimuli across the cortical hierarchy were both in line with previous findings and near-identical between age groups. That is, there was no age-related degradation in stimulus encoding in regions known to be critical for effective audiovisual spatial integration.

Chapter 5: Healthy Ageing Affects Audiovisual Speech Comprehension in Conditions of Asynchrony and Background Noise

Chapters 3 and 4 investigated audiovisual spatial integration—a form of multisensory interaction that has received limited considered in relation to ageing—and found that the age groups responded similarly both behaviourally and neurally. For the research in Chapter 5 we instead turned our attention to audiovisual speech integration, a function that is known to be affected by age. Specifically, we aimed to investigate the effects of age, audiovisual asynchrony, and background noise on sentence comprehension and synchrony judgement.

First, we confirmed behaviourally that sentence comprehension in both younger and older adults is impaired by the presence of masking babble noise and that this affects older adults more strongly, despite the groups being matched on comprehension in the absence of background noise. We also confirmed that the addition of a visual component improves comprehension of these stimuli to a similar degree in both age groups. Furthermore, we showed that asynchrony between the auditory and visual speech signals impaired comprehension in both age groups, an effect rarely measured in younger adults (see e.g. Grant, van Wassenhove, & Poeppel, 2004) and, to my knowledge, not previously demonstrated in the older population. Older adults' comprehension was, in fact, more impaired by stimulus asynchrony, an effect that coincided with an array of age differences in neural activation that included posterior right STS and the cognitive control/inhibition network.

In contrast, older adults performed as well as the younger group at detecting asynchrony in the audiovisual speech stimuli. This suggests that it is not awareness of asynchrony per se that is impaired with ageing, but the ability to inhibit and/or compensate for the incongruence. Similarly, the presence of background noise did not affect synchrony detection in either age group but did significantly impair comprehension, implying that these functions rely to some degree on separate mechanisms.

Unlike the results of the previous chapter, functional MRI data revealed age-related activation differences in both directions in response to these stimuli. Even in the absence of background noise, when the groups were thus behaviourally matched, there was a clear effect of older adults' impaired hearing leading to reduced auditory activations. More interestingly, older adults showed greatly increased responses to visual stimuli across multiple conditions. This was particularly true in the case of synchronous audiovisual stimuli presented with background noise. Taken together with the behavioural finding that both age groups showed similar amounts of multisensory benefit in this condition, we may conclude that older adults were extracting more information from the visual stimuli than the younger group.

Context, limitations, and future directions

Overall, our research fits well with prior investigations into the effects of ageing on multisensory integration and with cognitive ageing research more generally. Nothing here substantially contradicts previous findings; instead, our investigation of the ventriloquist effect, alongside novel applications of computational modelling and neuroimaging, provide new material for discussion and give clarification and mechanistic explanations for previously-described phenomena. Here I will outline some of the ways in which the results described in this thesis complement previous work, and discuss limitations and suggestions for future directions.

As I discuss in my review of multisensory ageing research (Chapter 1), in cases where responses to unisensory stimuli are compared with those to congruent multisensory stimuli, the degree of multisensory enhancement does not decline with age. Some studies, particularly those that use race model violations to measure enhancement of reaction times, even report greater enhancement in the older age group. I am unsure how robust this effect is, but there is certainly little evidence in the opposite direction. It is therefore reasonable to conclude that, under most conditions, multisensory enhancement is not impaired by the healthy ageing process. Our research supports this conclusion. Though we did not explicitly set out to test multisensory enhancement in the studies employing ventriloquist paradigms (Chapters 3 and 4), responses to congruent stimuli were similarly improved relative to unimodal auditory in both age groups (see, for example, Figure 4.1B). In the speech comprehension paradigm (Chapter 5) we did explicitly test the degree of multisensory enhancement and found it to be similar between age groups, or perhaps slightly greater in older adults. Future supplementary decoding analysis of our fMRI data could reveal how these enhancement effects are represented in the brain.

Our paradigms and analyses focused more on cases where younger and older participants must respond to some form of audiovisual incongruence, as these are the situations under which the greatest age differences have previously been shown. In Chapter 1, I identified a trend that multisensory stimuli with temporal conflict were the most likely to differentially affect younger and older adults. Our ventriloquist data support this finding through comparison with the now well-established phenomenon that the sound-induced flash illusion is stronger in older adults. As these two paradigms use similar low-level stimuli, the difference in behavioural outcomes must be attributed to the type of incongruence, i.e. temporal versus spatial. We have also established that older adults' responses to the ventriloquist task may be modelled using a BCI framework, and it is thus reasonable to consider age differences in other paradigms in the same way. For example, older adults' increased temporal binding in the sound-induced flash illusion may be attributed to a decrease in auditory reliability, an increase in visual reliability, or a larger common-source prior (Shams, Ma, & Beierholm, 2005). Though it is possible to speculate which of these is most likely, it would be more useful for this model to be fit to actual response data from older adults in future.

We have also, via our fMRI data, demonstrated a potential mechanism for older adults' impaired responses to incongruent multisensory stimuli. We showed that younger adults activated the cognitive control/inhibition network, often seen in response to e.g. Stroop stimuli, when presented with asynchronous speech stimuli in a challenging environment (background noise); older adults did not activate this network in the same way or to the same degree, and showed corresponding impairments of comprehension. Future response-contingent analyses on these data, that separate activation profiles based on the response given, could help us to better understand this effect. We must exercise caution in the degree to which we generalise this finding, however; as I have previously noted, multisensory integration is not a single function but a class of (often quite different) processes, and it would thus be more beneficial to carry out functional imaging investigations separately into the various behavioural effects that have been previously demonstrated. I again refer back to the frequently-replicated age differences in the sound-induced flash illusion, and recommend that the neural mechanisms underpinning this effect are investigated via fMRI. It must also be acknowledged that the two age groups tested in this study differed substantially from each other in their low-level sensory acuity (as measured by audiometry), which may have confounded some of the observed age differences in neural activations. This is, for example, the most probable explanation for overall higher auditory cortex activations in the younger adults. Future research may wish to match the age groups on their low-level sensory abilities, or to include results of audiometry as a covariate in the analysis.

It is worth briefly noting that in some our tasks there were small but obvious age differences that did not reach significance. These include the unimodal localisation responses in Chapter 3, bimodal localisation responses in Chapter 4, and degree of multisensory enhancement in Chapter 5. Clearly the absence of statistical significance does not necessary rule out the presence of a real-world effect, and it may be that some of the conclusions of these studies may have changed if we had used a different or larger set of participants. In the case of sound and bimodal localisation, these borderline findings fit well with an extensive study by Dobreva et al. (2012), who found that differences in localisation performance in older adults were present but small. However, where we do report age differences these are generally highly significant and often based on multiple converging analyses. Thus, replications of these studies may well reveal a greater number of significant age differences but are unlikely to find fewer.

The paradigms used in this thesis were exclusively audiovisual, and focused only on integration of spatial and speech information. However, I hope we have demonstrated that the approaches that have recently been applied to the understanding of all forms of multisensory integration in the young adult brain may also have a place in ageing research. For example, applying the paradigms of Ernst and colleagues to investigate the impact of ageing on visual-haptic cue weighting (Ernst & Banks, 2002; Ernst, 2007; Helbig & Ernst, 2008), or determining age differences in signatures of audiovisual weights using single-trial EEG analysis (Boyle, Kayser, and Kayser, 2017), may prove extremely informative. To more immediately follow the work described in this thesis, however, I would recommend a thorough behavioural investigation of forced-fusion weighting of audiovisual spatial cues, to rule out the presence of small age differences in audiovisual spatial integration. It would also be interesting to administer a speeded version of the ventriloquist paradigm that had sufficient conditions to allow fitting of the BCI model, and thus comparison of fitted parameters between speeded and non-speeded conditions in both age groups. This would directly support (or refute) our conclusion that older adults require more time to accumulate the necessary information before making an accurate response.

Conclusion

In conclusion, the work described in this thesis investigated the mechanisms underlying audiovisual integration in the young and healthy ageing brain. We used behavioural testing, psychophysics, computational modelling, and fMRI to characterise older adults' responses to multisensory location and speech stimuli. We found that older adults perform similarly, behaviourally and neurally, in response to audiovisual localisation tasks. In contrast, we showed substantial age differences in behavioural and neural responses to synchronous and asynchronous audiovisual speech stimuli presented with background noise. These findings have implications for future research into the effects of healthy ageing on multisensory integration.

REFERENCES

Acerbi, L., & Ma, W. J. (2017). Practical Bayesian Optimization for Model Fitting with Bayesian Adaptive Direct Search. *ArXiv:1705.04405 [q-Bio, Stat]*. Retrieved from http://arxiv.org/abs/1705.04405

Adam, R., & Noppeney, U. (2010). Prior auditory information shapes visual category-selectivity in ventral occipito-temporal cortex. *NeuroImage*, *52*(4), 1592–1602. doi:10.1016/j.neuroimage.2010.05.002

Adam, R., & Noppeney, U. (2014). A phonologically congruent sound boosts a visual target into perceptual awareness. *Frontiers in Integrative Neuroscience*, *8*. doi:10.3389/fnint.2014.00070

Aizenstein, H. J., Clark, K. A., Butters, M. A., Cochran, J., Stenger, V. A., Meltzer, C. C., … Carter, C. S. (2004). The BOLD Hemodynamic Response in Healthy Aging. *Journal of Cognitive Neuroscience*, *16*(5), 786–793. doi:10.1162/089892904970681

Alain, C., McDonald, K., & Van Roon, P. (2012). Effects of age and background noise on processing a mistuned harmonic in an otherwise periodic complex sound. *Hearing Research*, *283*(1–2), 126–135. doi:10.1016/j.heares.2011.10.007

Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, *14*(3), 257–262. doi:10.1016/j.cub.2004.01.029

Allard, R., Lagacé-Nadon, S., & Faubert, J. (2013). Feature tracking and aging. *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00427

Allison, L. K., Kiemel, T., & Jeka, J. J. (2018). Sensory-Challenge Balance Exercises Improve Multisensory Reweighting in Fall-Prone Older Adults. *Journal of Neurologic Physical Therapy : JNPT*, *42*(2), 84–93. doi:10.1097/NPT.0000000000000214

Allison, Leslie K., Kiemel, T., & Jeka, J. J. (2006). Multisensory reweighting of vision and touch is intact in healthy and fall-prone older adults. *Experimental Brain Research*, *175*(2), 342–352. doi:10.1007/s00221-006-0559-7

Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual Integration of Speech Falters under High Attention Demands. *Current Biology*, *15*(9), 839–843. doi:10.1016/j.cub.2005.03.046

Alsius, A., Navarra, J., & Soto-Faraco, S. (2007). Attention to touch weakens audiovisual speech integration. *Experimental Brain Research*, *183*(3), 399–404. doi:10.1007/s00221-007-1110-1

Altieri, N., Stevenson, R. A., Wallace, M. T., & Wenger, M. J. (2015). Learning to Associate Auditory and Visual Stimuli: Behavioral and Neural Mechanisms. *Brain Topography*, *28*(3), 479–493. doi:10.1007/s10548-013-0333-7

Ances, B. M., Liang, C. L., Leontiev, O., Perthen, J. E., Fleisher, A. S., Lansing, A. E., & Buxton, R. B. (2009). Effects of aging on cerebral blood flow, oxygen metabolism, and blood oxygenation level dependent responses to visual stimulation. *Human Brain Mapping*, *30*(4), 1120–1132. doi:10.1002/hbm.20574

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, *21*(3), 301–308. doi:10.1016/j.cogbrainres.2004.06.004

Anderson Gosselin, P., & Gagné, J.-P. (2011). Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. *International Journal of Audiology*, *50*(11), 786–792. doi:10.3109/14992027.2011.599870

Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences*, *110*(11), 4357–4362. doi:10.1073/pnas.1213555110

Andersson, J. L. R., Hutton, C., Ashburner, J., Turner, R., & Friston, K. (2001). Modeling Geometric Deformations in EPI Time Series. *NeuroImage*, *13*(5), 903–919. doi:10.1006/nimg.2001.0746

Andrés, P., Parmentier, F. B. R., & Escera, C. (2006). The effect of age on involuntary capture of attention by irrelevant sounds: A test of the frontal hypothesis of aging. *Neuropsychologia*, *44*(12), 2564–2568. doi:10.1016/j.neuropsychologia.2006.05.005

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, *26*(3), 839–851. doi:10.1016/j.neuroimage.2005.02.018

Audacity Team. (2018). Audacity (Version 2.2). Retrieved from https://www.audacityteam.org/

Avillac, M., Ben Hamed, S., & Duhamel, J.-R. (2007). Multisensory Integration in the Ventral Intraparietal Area of the Macaque Monkey. *Journal of Neuroscience*, *27*(8), 1922–1932. doi:10.1523/JNEUROSCI.2646-06.2007

Barraclough, N. E., Xiao, D., Baker, C. I., Oram, M. W., & Perrett, D. I. (2005). Integration of Visual and Auditory Information by Superior Temporal Sulcus Neurons Responsive to the Sight of Actions. *Journal of Cognitive Neuroscience*, *17*(3), 377–391. doi:10.1162/0898929053279586

Barrett, M. M., & Newell, F. N. (2015). Task-Specific, Age Related Effects in the Cross-Modal Identification and Localisation of Objects. *Multisensory Research*, *28*(1–2), 111–151. doi:10.1163/22134808-00002479

Barulli, D., & Stern, Y. (2013). Efficiency, capacity, compensation, maintenance, plasticity: emerging concepts in cognitive reserve. *Trends in Cognitive Sciences*, *17*(10), 502–509. doi:10.1016/j.tics.2013.08.012

Başkent, D., & Bazo, D. (2011). Audiovisual asynchrony detection and speech intelligibility in noise with moderate to severe sensorineural hearing impairment. *Ear and Hearing*, *32*(5), 582–592. doi:10.1097/AUD.0b013e31820fca23

Bates, S. L., & Wolbers, T. (2014). How cognitive aging affects multisensory integration of navigational cues. *Neurobiology of Aging*. doi:10.1016/j.neurobiolaging.2014.04.003

Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A*, *20*(7), 1391. doi:10.1364/JOSAA.20.001391

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), 809–823. doi:10.1016/S0896-6273(04)00070-4

Bedard, G., & Barnett-Cowan, M. (2015). Impaired timing of audiovisual events in the elderly. *Experimental Brain Research*, 1–10. doi:10.1007/s00221-015-4466-7

Bernstein, L. E., Auer, E. T., & Takayanagi, S. (2004). Auditory speech detection in noise enhanced by lipreading. *Speech Communication*, *44*(1–4), 5–18. doi:10.1016/j.specom.2004.10.011

Bertelson, P., Vroomen, J., De Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, *62*(2), 321–332. doi:10.3758/BF03205552

Blatter, D. D., Bigler, E. D., Gale, S. D., Johnson, S. C., Anderson, C. V., Burnett, B. M., … Horn, S. D. (1995). Quantitative volumetric analysis of brain MR: normative database spanning 5 decades of life. *American Journal of Neuroradiology*, *16*(2), 241–251.

Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer (Version 6.0.43). Retrieved from http://www.praat.org/

Botvinick, M., & Cohen, J. (1998). Rubber hands ‘feel’ touch that eyes see. *Nature*, *391*(6669), 756–756. doi:10.1038/35784

Boyle, S. C., Kayser, S. J., & Kayser, C. (2017). Neural correlates of multisensory reliability and perceptual weights emerge at early latencies during audio-visual integration. *The European Journal of Neuroscience, 46*(10), 2565–2577. doi:10.1111/ejn.13724

Braem, B., Honoré, J., Rousseaux, M., Saj, A., & Coello, Y. (2014). Integration of visual and haptic informations in the perception of the vertical in young and old healthy adults and right brain-damaged patients. *Neurophysiologie Clinique/Clinical Neurophysiology*, *44*(1), 41–48. doi:10.1016/j.neucli.2013.10.137

Bresciani, J.-P., Dammeier, F., & Ernst, M. O. (2008). Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. *Brain Research Bulletin*, *75*(6), 753–760. doi:10.1016/j.brainresbull.2008.01.009

Brewer, A. A., & Barton, B. (2016). Maps of the Auditory Cortex. *Annual Review of Neuroscience*, *39*(1), 385–407. doi:10.1146/annurev-neuro-070815-014045

Brody, H. (1955). Organization of the cerebral cortex. III. A study of aging in the human cerebral cortex. *Journal of Comparative Neurology*, *102*(2), 511–556. doi:10.1002/cne.901020206

Brooks, C. J., Anderson, A. J., Roach, N. W., McGraw, P. V., & McKendrick, A. M. (2015). Age-related changes in auditory and visual interactions in temporal rate perception. *Journal of Vision*, *15*(16), 2. doi:10.1167/15.16.2

Bugg, J. M., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2007). Age Differences in Stroop Interference: Contributions of General Slowing and Task-Specific Deficits. *Aging, Neuropsychology, and Cognition*, *14*(2), 155–167. doi:10.1080/138255891007065

Burr, D., Banks, M. S., & Morrone, M. C. (2009). Auditory dominance over vision in the perception of interval duration. *Experimental Brain Research*, *198*(1), 49. doi:10.1007/s00221-009-1933-z

Burr, D., & Gori, M. (2012). Multisensory Integration Develops Late in Humans. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press/Taylor & Francis. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK92864/

Busse, L., Roberts, K. C., Crist, R. E., Weissman, D. H., & Woldorff, M. G. (2005). The spread of attention across modalities and space in a multisensory object. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(51), 18751–18756. doi:10.1073/pnas.0507704102

Butler, J. S., Smith, S. T., Campos, J. L., & Bülthoff, H. H. (2010). Bayesian integration of visual and vestibular signals for heading. *Journal of Vision*, *10*(11), 23–23. doi:10.1167/10.11.23

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and Aging*, *17*(1), 85–100. doi:10.1037/0882-7974.17.1.85

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging Gracefully: Compensatory Brain Activity in High-Performing Older Adults. *NeuroImage*, *17*(3), 1394–1402. doi:10.1006/nimg.2002.1280

Callejas, A., Lupiáñez, J., & Tudela, P. (2004). The three attentional networks: On their independence and interactions. *Brain and Cognition*, *54*(3), 225–227. doi:10.1016/j.bandc.2004.02.012

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., … David, A. S. (1997). Activation of Auditory Cortex During Silent Lipreading. *Science, 276*(5312), 593–596. doi:10.1126/science.276.5312.593

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*, *10*(11), 649–657. doi:10.1016/S0960-9822(00)00513-3

Campbell, C. S., & Massaro, D. W. (1997). Perception of Visible Speech: Influence of Spatial Quantization. *Perception*, *26*(5), 627–644. doi:10.1068/p260627

Campbell, K. L., Al-Aidroos, N., Fatt, R., Pratt, J., & Hasher, L. (2010). The effects of multisensory targets on saccadic trajectory deviations: eliminating age differences. *Experimental Brain Research*, *201*(3), 385–392. doi:10.1007/s00221-009-2045-5

Campbell, R. (2008). The processing of audio-visual speech: empirical and neural bases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 1001–1010. doi:10.1098/rstb.2007.2155

Campbell, Ruth, & Dodd, B. (1980). Hearing by eye. *Quarterly Journal of Experimental Psychology*, *32*(1), 85–99. doi:10.1080/00335558008248235

Chadwick, M. J., Bonnici, H. M., & Maguire, E. A. (2014). CA3 size predicts the precision of memory recall. *Proceedings of the National Academy of Sciences*, *111*(29), 10720–10725. doi:10.1073/pnas.1319641111

Chan, J. S., Connolly, S. K., & Setti, A. (2018). The Number of Stimulus-Onset Asynchronies Affects the Perception of the Sound-Induced Flash Illusion in Young and Older Adults. *Multisensory Research*, *31*(3–4), 175–190. doi:10.1163/22134808-00002605

Chan, J. S., Maguinness, C., Lisiecka, D., Setti, A., & Newell, F. N. (2012). Evidence for Crossmodal Interactions across Depth on Target Localisation Performance in a Spatial Array. *Perception*, *41*(7), 757–773. doi:10.1068/p7230

Chan, Y. M., Pianta, M. J., & McKendrick, A. M. (2014a). Older age results in difficulties separating auditory and visual signals in time. *Journal of Vision*, *14*(11). doi:10.1167/14.11.13

Chan, Y. M., Pianta, M. J., & McKendrick, A. M. (2014b). Reduced audiovisual recalibration in the elderly. *Frontiers in Aging Neuroscience*, *6*. doi:10.3389/fnagi.2014.00226

Cherry, B. J., Adamson, M., Duclos, A., & Hellige, J. B. (2005). Aging and Individual Variation in Interhemispheric Collaboration and Hemispheric Asymmetry. *Aging, Neuropsychology, and Cognition*, *12*(4), 316–339. doi:10.1080/17444128.2005.10367004

Churchill, J., Stanis, J., Press, C., Kushelev, M., & Greenough, W. T. (2003). Is procedural memory relatively spared from age effects? *Neurobiology of Aging*, *24*(6), 883–892. doi:10.1016/S0197-4580(02)00194-X

Cienkowski, K. M., & Carney, A. E. (2002). Auditory-visual speech perception and aging. *Ear and Hearing*, *23*(5), 439–449.

Cohen, G. (2014, October 10). The Effects of Aging on Autobiographical Memory. doi:10.4324/9781315784250-7

Colcombe, S. J., Erickson, K. I., Scalf, P. E., Kim, J. S., Prakash, R., McAuley, E., … Kramer, A. F. (2006). Aerobic Exercise Training Increases Brain Volume in Aging Humans. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *61*(11), 1166–1170.

Colin, C., Radeau, M., & Deltenre, P. (2005). Top-down and bottom-up modulation of audiovisual integration in speech. *European Journal of Cognitive Psychology*, *17*(4), 541–560. doi:10.1080/09541440440000168

Colonius, H. (1988). Modeling the redundant signals effect by specifying the hazard function. *Perception & Psychophysics*, *43*(6), 604–606. doi:10.3758/BF03207750

Colonius, H., & Diederich, A. (2011). Computing an optimal time window of audiovisual integration in focused attention tasks: illustrated by studies on effect of age and prior knowledge. *Experimental Brain Research*, *212*(3), 327–337. doi:10.1007/s00221-011-2732-x

Conrey, B., & Pisoni, D. B. (2006). Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *The Journal of the Acoustical Society of America*, *119*(6), 4065–4073. doi:10.1121/1.2195091

Courtney, S. M., Petit, L., Maisog, J. M., Ungerleider, L. G., & Haxby, J. V. (1998). An Area Specialized for Spatial Working Memory in Human Frontal Cortex. *Science*, *279*(5355), 1347–1351. doi:10.1126/science.279.5355.1347

Cowan, N., & Barron, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics*, *41*(5), 393–401. doi:10.3758/BF03203031

Crivello, F., Tzourio-Mazoyer, N., Tzourio, C., & Mazoyer, B. (2014). Longitudinal assessment of global and regional rate of grey matter atrophy in 1,172 healthy older adults: modulation by sex and age. *PloS One*, *9*(12), e114478. doi:10.1371/journal.pone.0114478

Dalton, P., Doolittle, N., Nagata, H., & Breslin, P. A. S. (2000). The merging of the senses: integration of subthreshold taste and smell. *Nature Neuroscience*, *3*(5), 431–432. doi:10.1038/74797

Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2008). Que PASA? The posterior-anterior shift in aging. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(5), 1201–1209. doi:10.1093/cercor/bhm155

de Boer-Schellekens, L., & Vroomen, J. (2013). Multisensory integration compensates loss of sensitivity of visual temporal order in the elderly. *Experimental Brain Research*, *232*(1), 253–262. doi:10.1007/s00221-013-3736-5

de Dieuleveult, A. L., Siemonsma, P. C., van Erp, J. B. F., & Brouwer, A.-M. (2017). Effects of Aging in Multisensory Integration: A Systematic Review. *Frontiers in Aging Neuroscience*, *9*. doi:10.3389/fnagi.2017.00080

de Fockert, J. W., Ramchurn, A., van Velzen, J., Bergström, Z., & Bunce, D. (2009). Behavioral and ERP evidence of greater distractor processing in old age. *Brain Research*, *1282*(Supplement C), 67–73. doi:10.1016/j.brainres.2009.05.060

DeLoss, D., & Andersen, G. (2015). Age-Related Differences in the Flash-Induced Beep Illusion. *Journal of Vision*, *15*(12), 849. doi:10.1167/15.12.849

DeLoss, D. J., & Andersen, G. J. (2015). Aging, Spatial Disparity, and the Sound-Induced Flash Illusion. *PLOS ONE*, *10*(11), e0143773. doi:10.1371/journal.pone.0143773

DeLoss, D. J., Pierce, R. S., & Andersen, G. J. (2013). Multisensory Integration, Aging, and the Sound-Induced Flash Illusion. *Psychology and Aging*, *28*(3), 802–812. doi:10.1037/a0033289

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health and Lifestyle Survey. *Psychology and Aging*, *21*(1), 62–73. doi:10.1037/0882-7974.21.1.62

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*(1), 193–222. doi:10.1146/annurev.ne.18.030195.001205

D'Esposito, M., Deouell, L. Y., & Gazzaley, A. (2003). Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging. *Nature Reviews Neuroscience*, *4*(11), 863–872. doi:10.1038/nrn1246

D'Esposito, M., Zarahn, E., Aguirre, G. K., & Rypma, B. (1999). The Effect of Normal Aging on the Coupling of Neural Activity to the Bold Hemodynamic Response. *NeuroImage*, *10*(1), 6–14. doi:10.1006/nimg.1999.0444

Dey, A., & Sommers, M. S. (2015). Age-related Differences in Inhibitory Control Predict Audiovisual Speech Perception. *Psychology and Aging*, *30*(3), 634–646. doi:10.1037/pag0000033

Diaconescu, A. O., Hasher, L., & McIntosh, A. R. (2013). Visual dominance and multisensory integration changes with age. *NeuroImage*, *65*, 152–166. doi:10.1016/j.neuroimage.2012.09.057

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, *66*(8), 1388–1404. doi:10.3758/BF03195006

Diederich, A., Colonius, H., & Schomburg, A. (2008). Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia*, *46*(10), 2556–2562. doi:10.1016/j.neuropsychologia.2008.03.026

Dobreva, M. S., O'Neill, W. E., & Paige, G. D. (2012). Influence of age, spatial memory, and ocular fixation on localization of auditory, visual, and bimodal targets by human subjects. *Experimental Brain Research*, *223*(4), 441–455. doi:10.1007/s00221-012-3270-x

Donohue, S. E., Green, J. J., & Woldorff, M. G. (2015). The effects of attention on the temporal integration of multisensory stimuli. *Frontiers in Integrative Neuroscience*, *9*. doi:10.3389/fnint.2015.00032

Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *The Journal of the Acoustical Society of America*, *76*(1), 87–96. doi:10.1121/1.391011

Ebaid, D., Crewther, S. G., MacCalman, K., Brown, A., & Crewther, D. P. (2017). Cognitive Processing Speed across the Lifespan: Beyond the Influence of Motor Speed. *Frontiers in Aging Neuroscience*, *9*. doi:10.3389/fnagi.2017.00062

Eckert, M. A., Keren, N. I., Roberts, D. R., Calhoun, V. D., & Harris, K. C. (2010). Age-Related Changes in Processing Speed: Unique Contributions of Cerebellar and Prefrontal Cortex. *Frontiers in Human Neuroscience*, *4*. doi:10.3389/neuro.09.010.2010

Edmister, W. B., Talavage, T. M., Ledden, P. J., & Weisskoff, R. M. (1999). Improved auditory cortex imaging using clustered volume acquisitions. *Human Brain Mapping*, *7*(2), 89–97. doi:10.1002/(SICI)1097-0193(1999)7:2<89::AID-HBM2>3.0.CO;2-N

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335. doi:10.1016/j.neuroimage.2004.12.034

Elliott, M. T., Wing, A. M., & Welchman, A. E. (2011). The effect of ageing on multisensory integration for the control of movement timing. *Experimental Brain Research*, *213*(2–3), 291–298. doi:10.1007/s00221-011-2740-x

Elliott, M. T., Wing, A. M., & Welchman, A. E. (2014). Moving in time: Bayesian causal inference explains movement coordination to auditory beats. *Proceedings of the Royal Society of London B: Biological Sciences*, *281*(1786), 20140751. doi:10.1098/rspb.2014.0751

Erber, N. P. (1975). Auditory-Visual Perception of Speech. *Journal of Speech and Hearing Disorders*, *40*(4), 481. doi:10.1044/jshd.4004.481

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. doi:10.3758/BF03203267

Eriksen, C. W., Hamlin, R. M., & Breitmeyer, R. G. (1970). Temporal factors in visual perception as related to aging. *Perception & Psychophysics*, *7*(6), 354–356. doi:10.3758/BF03208664

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, *7*(5), 7–7. doi:10.1167/7.5.7

Ernst, Marc O. (2006). A Bayesian view on multimodal cue integration. *Human Body Perception from the inside Out*, 105–131.

Ernst, Marc O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. doi:10.1038/415429a

Esiri, M. (2007). Ageing and the brain. *The Journal of Pathology*, *211*(2), 181–187. doi:10.1002/path.2089

Fan, J, McCandliss, B. D., Fossella, J., Flombaum, J., & Posner, M. (2005). The activation of attentional networks. *NeuroImage*, *26*(2), 471–479. doi:10.1016/j.neuroimage.2005.02.004

Fan, Jin, McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the Efficiency and Independence of Attentional Networks. *Journal of Cognitive Neuroscience*, *14*(3), 340–347. doi:10.1162/089892902317361886

Faubert, J. (2002). Visual perception and aging. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *56*(3), 164–176. doi:10.1037/h0087394

Fera, F., Weickert, T. W., Goldberg, T. E., Tessitore, A., Hariri, A., Das, S., … Mattay, V. S. (2005). Neural Mechanisms Underlying Probabilistic Category Learning in Normal Aging. *Journal of Neuroscience*, *25*(49), 11340–11348. doi:10.1523/JNEUROSCI.2736-05.2005

Fetsch, C. R., DeAngelis, G. C., & Angelaki, D. E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience*, *14*(6), 429–442. doi:10.1038/nrn3503

Fetsch, C. R., Pouget, A., DeAngelis, G. C., & Angelaki, D. E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. *Nature Neuroscience*, *15*(1), 146–154. doi:10.1038/nn.2983

Fetsch, C. R., Turner, A. H., DeAngelis, G. C., & Angelaki, D. E. (2009). Dynamic Reweighting of Visual and Vestibular Cues during Self-Motion Perception. *The Journal of Neuroscience*, *29*(49), 15601–15612. doi:10.1523/JNEUROSCI.2574-09.2009

Fiacconi, C. M., Harvey, E. C., Sekuler, A. B., & Bennett, P. J. (2013). The Influence of Aging on Audiovisual Temporal Order Judgments. *Experimental Aging Research*, *39*(2), 179–193. doi:10.1080/0361073X.2013.761896

Fischl, B. (2012). FreeSurfer. *NeuroImage*, *62*(2), 774–781. doi:10.1016/j.neuroimage.2012.01.021

FitzGerald, T. H. B., Friston, K. J., & Dolan, R. J. (2012). Action-Specific Value Signals in Reward-Related Regions of the Human Brain. *Journal of Neuroscience*, *32*(46), 16417–16423. doi:10.1523/JNEUROSCI.3254-12.2012

Fitzgibbons, P. J., & Gordon-Salant, S. (2004). Age effects on discrimination of timing in auditory sequences. *The Journal of the Acoustical Society of America*, *116*(2), 1126–1134. doi:10.1121/1.1765192

Fjell, A. M., & Walhovd, K. B. (2010). Structural Brain Changes in Aging: Courses, Causes and Cognitive Consequences. *Reviews in the Neurosciences*, *21*(3). doi:10.1515/REVNEURO.2010.21.3.187

Fjell, A. M., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., … Walhovd, K. B. (2009). High Consistency of Regional Cortical Thinning in Aging across Multiple Samples. *Cerebral Cortex*, *19*(9), 2001–2012. doi:10.1093/cercor/bhn232

Forster, B., Cavina-Pratesi, C., Aglioti, S. M., & Berlucchi, G. (2002). Redundant target effect and intersensory facilitation from visual-tactile interactions in simple reaction time. *Experimental Brain Research*, *143*(4), 480–487. doi:10.1007/s00221-002-1017-9

Forstmann, B. U., Tittgemeyer, M., Wagenmakers, E.-J., Derrfuss, J., Imperati, D., & Brown, S. (2011). The Speed-Accuracy Tradeoff in the Elderly Brain: A Structural Model-Based Approach. *Journal of Neuroscience*, *31*(47), 17242–17249. doi:10.1523/JNEUROSCI.0309-11.2011

Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., … Murray, M. M. (2002). Auditory-Somatosensory Multisensory Processing in Auditory Association

Cortex: An fMRI Study. *Journal of Neurophysiology*, *88*(1), 540–543. doi:10.1152/jn.2002.88.1.540

Fozard, J. L., Vercruyssen, M., Reynolds, S. L., Hancock, P. A., & Quilter, R. E. (1994). Age Differences and Changes in Reaction Time: The Baltimore Longitudinal Study of Aging. *Journal of Gerontology*, *49*(4), P179–P189. doi:10.1093/geronj/49.4.P179

Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, *147*(3), 332–343. doi:10.1007/s00221-002-1262-y

Freiherr, J., Lundström, J. N., Habel, U., & Reetz, K. (2013). Multisensory integration mechanisms during aging. *Frontiers in Human Neuroscience*, *7*, 863. doi:10.3389/fnhum.2013.00863

Frisina, R. D., & Walton, J. P. (2006). Age-related structural and functional changes in the cochlear nucleus. *Hearing Research*, *216–217*, 216–223. doi:10.1016/j.heares.2006.02.003

Friston, K., Chu, C., Mourão-Miranda, J., Hulme, O., Rees, G., Penny, W., & Ashburner, J. (2008). Bayesian decoding of brain images. *NeuroImage*, *39*(1), 181–205. doi:10.1016/j.neuroimage.2007.08.013

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), 189–210. doi:10.1002/hbm.460020402

Friston, Karl J., Penny, W. D., & Glaser, D. E. (2005). Conjunction revisited. *NeuroImage*, *25*(3), 661–667. doi:10.1016/j.neuroimage.2005.01.013

Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-Related fMRI: Characterizing Differential Responses. *NeuroImage*, *7*(1), 30–40. doi:10.1006/nimg.1997.0306

Füllgrabe, C., Moore, B. C. J., & Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience*, *6*. doi:10.3389/fnagi.2014.00347

Gates, G. A., & Mills, J. H. (2005). Presbycusis. *The Lancet*, *366*(9491), 1111–1120. doi:10.1016/S0140-6736(05)67423-5

Gau, R., & Noppeney, U. (2016). How prior expectations shape multisensory perception. *NeuroImage*, *124*, *Part A*, 876–886. doi:10.1016/j.neuroimage.2015.09.045

Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, *8*(10), 1298–1300. doi:10.1038/nn1543

Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory Integration of Dynamic Faces and Voices in Rhesus Monkey Auditory Cortex. *Journal of Neuroscience*, *25*(20), 5004–5012. doi:10.1523/JNEUROSCI.0799-05.2005

Ghazanfar, A., & Schroeder, C. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*(6), 278–285. doi:10.1016/j.tics.2006.04.008

Giard, M. H., & Peronnet, F. (1999). Auditory-Visual Integration during Multimodal Object Recognition in Humans: A Behavioral and Electrophysiological Study. *Journal of Cognitive Neuroscience*, *11*(5), 473–490. doi:10.1162/089892999563544

Gillmeister, H., & Eimer, M. (2007). Tactile enhancement of auditory detection and perceived loudness. *Brain Research*, *1160*, 58–68. doi:10.1016/j.brainres.2007.03.041

Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S., De Stefano, N., & Johansen-Berg, H. (2010). Age-related changes in grey and white matter structure throughout adulthood. *NeuroImage*, *51*(3), 943–951. doi:10.1016/j.neuroimage.2010.03.004

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N. A., Friston, K. J., & Frackowiak, R. S. J. (2001). A Voxel-Based Morphometric Study of Ageing in 465 Normal Adult Human Brains. *NeuroImage*, 21–36.

Gordon, B. A., Rykhlevskaia, E. I., Brumback, C. R., Lee, Y., Elavsky, S., Konopack, J. F., … Fabiani, M. (2008). Neuroanatomical correlates of aging, cardiopulmonary fitness level, and education. *Psychophysiology*, *45*(5), 825–838. doi:10.1111/j.1469-8986.2008.00676.x

Gordon, M. S., & Allen, S. (2009). Audiovisual Speech in Older and Younger Adults: Integrating a Distorted Visual Signal With Speech in Noise. *Experimental Aging Research*, *35*(2), 202–219. doi:10.1080/03610730902720398

Grady, C, & Craik, F. I. M. (2000). Changes in memory processing with age. *Current Opinion in Neurobiology*, *10*(2), 224–231. doi:10.1016/S0959-4388(00)00073-8

Grady, C. L., & Garrett, D. D. (2014). Understanding variability in the BOLD signal and why it matters for aging. *Brain Imaging and Behavior*, *8*(2), 274–283. doi:10.1007/s11682-013-9253-0

Grady, Cheryl. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience*, *13*(7), 491. doi:10.1038/nrn3256

Grant, K. W., & Greenberg, S. (2001). Speech intelligibility derived from asynchronous processing of auditory-visual information. In *AVSP 2001-International Conference on Auditory-Visual Speech Processing*. Retrieved from http://www.isca-speech.org/archive_open/avsp01/av01_132.html

Grant, K. W., & Seitz, P. F. (1998). Measures of auditory–visual integration in nonsense syllables and sentences. *The Journal of the Acoustical Society of America*, *104*(4), 2438–2450. doi:10.1121/1.423751

Grant, K. W., & Seitz, P.-F. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *The Journal of the Acoustical Society of America*, *108*(3), 1197. doi:10.1121/1.1288668

Grant, K. W., van Wassenhove, V., & Poeppel, D. (2004). Detection of auditory (cross-spectral) and auditory–visual (cross-modal) synchrony. *Speech Communication*, *44*(1–4), 43–53. doi:10.1016/j.specom.2004.06.004

Grant, K. W., Walden, B. E., & Seitz, P. F. (1998). Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *The Journal of the Acoustical Society of America*, *103*(5), 2677–2690. doi:10.1121/1.422788

Grill-Spector, K., & Malach, R. (2004). The Human Visual Cortex. *Annual Review of Neuroscience*, *27*(1), 649–677. doi:10.1146/annurev.neuro.27.070203.144220

Guerreiro, M. J. S., Adam, J. J., Gerven, V., & M, P. W. (2012). Automatic Selective Attention as a Function of Sensory Modality in Aging. *The Journals of Gerontology: Series B*, *67B*(2), 194–202. doi:10.1093/geronb/gbr090

Guerreiro, M. J. S., Anguera, J. A., Mishra, J., Van Gerven, P. W. M., & Gazzaley, A. (2014). Age-equivalent Top–Down Modulation during Cross-modal Selective Attention. *Journal of Cognitive Neuroscience*, *26*(12), 2827–2839. doi:10.1162/jocn_a_00685

Guerreiro, M. J. S., Eck, J., Moerel, M., Evers, E. A. T., & Gerven, P. W. M. V. (2015). Top-down modulation of visual and auditory cortical processing in aging. *Behavioural Brain Research*, *278*(Supplement C), 226–234. doi:10.1016/j.bbr.2014.09.049

Guerreiro, M. J. S., Murphy, D. R., & Van Gerven, P. W. M. (2013). Making sense of age-related distractibility: The critical role of sensory modality. *Acta Psychologica*, *142*(2), 184–194. doi:10.1016/j.actpsy.2012.11.007

Guerreiro, M. J. S., & Van Gerven, P. W. M. (2011). Now you see it, now you don't: Evidence for age-dependent and age-independent cross-modal distraction. *Psychology and Aging*, *26*(2), 415–426. doi:10.1037/a0021507

Guerreiro, M., Murphy, D., & Van Gerven, P. (2010). *The Role of Sensory Modality in Age-Related Distraction: A Critical Review and a Renewed View* (Vol. 136). doi:10.1037/a0020731

Gunning-Dixon, F. M., Brickman, A. M., Cheng, J. C., & Alexopoulos, G. S. (2009). Aging of cerebral white matter: a review of MRI findings. *International Journal of Geriatric Psychiatry*, *24*(2), 109–117. doi:10.1002/gps.2087

Gunning-Dixon, F. M., & Raz, N. (2003). Neuroanatomical correlates of selected executive functions in middle-aged and older adults: a prospective MRI study. *Neuropsychologia*, *41*(14), 1929–1941. doi:10.1016/S0028-3932(03)00129-5

Guttmann, C. R., Jolesz, F. A., Kikinis, R., Killiany, R. J., Moss, M. B., Sandor, T., & Albert, M. S. (1998). White matter changes with normal aging. *Neurology*, *50*(4), 972–978.

Hairston, W. D., Laurienti, P. J., Mishra, G., Burdette, J. H., & Wallace, M. T. (2003). Multisensory enhancement of localization under conditions of induced myopia. *Experimental Brain Research*, *152*(3), 404–408. doi:10.1007/s00221-003-1646-7

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., … Bowtell, R. W. (1999). ?sparse? temporal sampling in auditory fMRI. *Human Brain Mapping*, *7*(3), 213–223. doi:10.1002/(SICI)1097-0193(1999)7:3<213::AID-HBM5>3.0.CO;2-N

Harrington, L. K., & Peck, C. K. (1998). Spatial disparity affects visual-auditory interactions in human sensorimotor processing. *Experimental Brain Research*, *122*(2), 247–252. doi:10.1007/s002210050512

Hasher, L., & Zacks, R. T. (1988). Working Memory, Comprehension, and Aging: A Review and a New View. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 22, pp. 193–225). Academic Press. doi:10.1016/S0079-7421(08)60041-9

Head, D., Kennedy, K. M., Rodrigue, K. M., & Raz, N. (2009). Age differences in perseveration: Cognitive and neuroanatomical mediators of performance on the Wisconsin Card Sorting Test. *Neuropsychologia*, *47*(4), 1200–1203. doi:10.1016/j.neuropsychologia.2009.01.003

Hebart, M. N., Görgen, K., & Haynes, J.-D. (2015). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Frontiers in Neuroinformatics*, *8*, 88. doi:10.3389/fninf.2014.00088

Helbig, H. B., & Ernst, M. O. (2008). Visual-haptic cue weighting is independent of modality-specific attention. *Journal of Vision*, *8*(1), 21–21. doi:10.1167/8.1.21

Helfer, K. S. (1998). Auditory and Auditory-Visual Recognition of Clear and Conversational Speech by Older Adults. *Journal of the American Academy of Audiology*, *9*(3), 9.

Helfer, K. S., & Freyman, R. L. (2008). Aging and Speech-on-Speech Masking. *Ear and Hearing*, *29*(1), 87–98. doi:10.1097/AUD.0b013e31815d638b

Herbert, A. M., Overbury, O., Singh, J., & Faubert, J. (2002). Aging and Bilateral Symmetry Detection. *The Journals of Gerontology: Series B*, *57*(3), P241–P245. doi:10.1093/geronb/57.3.P241

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, *63*(3), 289–293.

Hertzog, C., Vernon, M. C., & Rypma, B. (1993). Age Differences in Mental Rotation Task Performance: The Influence of Speed/Accuracy Tradeoffs. *Journal of Gerontology*, *48*(3), P150–P156. doi:10.1093/geronj/48.3.P150

Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart: exercise effects on brain and cognition. *Nature Reviews Neuroscience*, *9*(1), 58–65.

Hopkins, K., & Moore, B. C. J. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *The Journal of the Acoustical Society of America*, *130*(1), 334–349. doi:10.1121/1.3585848

Hugenschmidt, C. E. a, Mozolic, J. L. a, & Laurienti, P. J. (2009). Suppression of multisensory integration by modality-specific attention in aging. *Neuroreport*, *20*(4), 349–353. doi:10.1097/WNR.0b013e328323ab07

Hugenschmidt, C. E., Peiffer, A. M., McCoy, T. P., Hayasaka, S., & Laurienti, P. J. (2009). Preservation of crossmodal selective attention in healthy aging. *Experimental Brain Research*, *198*(2–3), 273–285. doi:10.1007/s00221-009-1816-3

Hulme, O. J., Skov, M., Chadwick, M. J., Siebner, H. R., & Ramsøy, T. Z. (2014). Sparse encoding of automatic visual association in hippocampal networks. *NeuroImage*, *102*, 458–464. doi:10.1016/j.neuroimage.2014.07.020

Humes, L. E. (2002). Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *The Journal of the Acoustical Society of America*, *112*(3), 1112–1132. doi:10.1121/1.1499132

Humes, L. E. (2007). The Contributions of Audibility and Cognitive Factors to the Benefit Provided by Amplified Speech to Older Adults. *Journal of the American Academy of Audiology*, *18*(7), 590–603. doi:10.3766/jaaa.18.7.6

Huyse, A., Leybaert, J., & Berthommier, F. (2014). Effects of aging on audio-visual speech integration. *The Journal of the Acoustical Society of America*, *136*(4), 1918–1931. doi:10.1121/1.4894685

Hwang, J.-H., Li, C.-W., Wu, C.-W., Chen, J.-H., & Liu, T.-C. (2007). Aging Effects on the Activation of the Auditory Cortex during Binaural Speech Listening in White Noise: An fMRI Study. *Audiology and Neurotology*, *12*(5), 285–294. doi:10.1159/000103209

Ibrahim, L. A., Mesik, L., Ji, X., Fang, Q., Li, H., Li, Y., … Tao, H. W. (2016). Cross-Modality Sharpening of Visual Cortical Processing through Layer-1-Mediated Inhibition and Disinhibition. *Neuron*, *89*(5), 1031–1045. doi:10.1016/j.neuron.2016.01.027

Jacobs, B., Driscoll, L., & Schall, M. (1997). Life-span dendritic and spine changes in areas 10 and 18 of human cortex: A quantitative golgi study. *The Journal of Comparative Neurology*, *386*(4), 661–680. doi:10.1002/(SICI)1096-9861(19971006)386:4<661::AID-CNE11>3.0.CO;2-N

Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vision Research*, *39*(21), 3621–3629. doi:10.1016/S0042-6989(99)00088-7

Jernigan, T. L., Archibald, S. L., Berhow, M. T., Sowell, E. R., Foster, D. S., & Hesselink, J. R. (1991). Cerebral structure on MRI, Part I: Localization of age-related changes. *Biological Psychiatry*, *29*(1), 55–67. doi:10.1016/0006-3223(91)90210-D

Jernigan, T. L., Archibald, S. L., Fennema-Notestine, C., Gamst, A. C., Stout, J. C., Bonner, J., & Hesselink, J. R. (2001). Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiology of Aging*, *22*(4), 581–594. doi:10.1016/S0197-4580(01)00217-2

Jorm, A. F., Rodgers, B., Henderson, A. S., Korten, A. E., Jacomb, P. A., Christensen, H., & Mackinnon, A. (1998). Occupation type as a predictor of cognitive decline and dementia in old age. *Age and Ageing, 27*(4), 477–483. doi: 10.1093/ageing/27.4.477

Kalaria, R. N. (2010). Vascular basis for brain degeneration: faltering controls and risk factors for dementia. *Nutrition Reviews*, *68*(suppl_2), S74–S87. doi:10.1111/j.1753-4887.2010.00352.x

Kannurpatti, S. S., & Biswal, B. B. (2008). Detection and scaling of task-induced fMRI-BOLD response using resting state fluctuations. *NeuroImage*, *40*(4), 1567–1574. doi:10.1016/j.neuroimage.2007.09.040

Kennedy, K. M., Rodrigue, K. M., Head, D., Gunning-Dixon, F., & Raz, N. (2009). Neuroanatomical and cognitive mediators of age-related differences in perceptual priming and learning. *Neuropsychology*, *23*(4), 475–491. doi:10.1037/a0015377

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *116*(4), 2395–2405. doi:10.1121/1.1784440

Klein, R., Klein, B. E., Linton, K. L., & De Mets, D. L. (1991). The Beaver Dam Eye Study: visual acuity. *Ophthalmology*, *98*(8), 1310–1315.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? In *30th European Conference on Visual Perception*.

Kochunov, P., Coyle, T., Lancaster, J., Robin, D. A., Hardies, J., Kochunov, V., … Fox, P. T. (2010). Processing speed is correlated with cerebral health markers in the frontal lobes as quantified by neuroimaging. *NeuroImage*, *49*(2), 1190–1199. doi:10.1016/j.neuroimage.2009.09.052

Kohama, S. G., Rosene, D. L., & Sherman, L. S. (2012). Age-related changes in human and non-human primate white matter: from myelination disturbances to cognitive decline. *AGE*, *34*(5), 1093–1110. doi:10.1007/s11357-011-9357-7

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal Inference in Multisensory Perception. *PLoS ONE*, *2*(9), e943. doi:10.1371/journal.pone.0000943

Kranczioch, C., Knäpper, W., & Klein, F. (2018). Examining the effect of age on auditory enhancement of visual-temporal attention. *Archives of Psychology*, *2*(3). Retrieved from http://www.archivesofpsychology.org/index.php/aop/article/view/43

Lachs, L., & Pisoni, D. B. (2004). Specification of cross-modal source information in isolated kinematic displays of speech. *The Journal of the Acoustical Society of America*, *116*(1), 507–518. doi:10.1121/1.1757454

Langenecker, S. A., Nielson, K. A., & Rao, S. M. (2004). fMRI of healthy older adults during Stroop interference. *NeuroImage*, *21*(1), 192–200. doi:10.1016/j.neuroimage.2003.08.027

Laurienti, P. J., Burdette, J. H., Maldjian, J. A., & Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiology of Aging*, *27*(8), 1155–1163. doi:10.1016/j.neurobiolaging.2005.05.024

Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H., & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research*, *158*(4), 405–414. doi:10.1007/s00221-004-1913-2

Lee, H., & Noppeney, U. (2014). Temporal prediction errors in visual and auditory cortices. *Current Biology*, *24*(8), R309–R310. doi:10.1016/j.cub.2014.02.007

Legault, I., Gagné, J.-P., Rhoualem, W., & Anderson-Gosselin, P. (2010). The effects of blurred vision on auditory-visual speech perception in younger and older adults. *International Journal of Audiology*, *49*(12), 904–911. doi:10.3109/14992027.2010.509112

Lehmann, C., Herdener, M., Esposito, F., Hubl, D., di Salle, F., Scheffler, K., … Seifritz, E. (2006). Differential patterns of multisensory interactions in core and belt areas of human auditory cortex. *NeuroImage*, *31*(1), 294–300. doi:10.1016/j.neuroimage.2005.12.038

Lewald, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cognitive Brain Research*, *16*(3), 468–478. doi:10.1016/S0926-6410(03)00074-0

Lewkowicz, D. J. (1996). Perception of auditory–visual temporal synchrony in human infants. *Journal of Experimental Psychology: Human Perception and Performance, 22*(5), 1094–1106. doi:10.1037/0096-1523.22.5.1094

Lewkowicz, D. J. (2010). Infant perception of audio-visual speech synchrony. *Developmental Psychology, 46*(1), 66–77. doi:10.1037/a0015579

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging*, *9*(3), 339–355. doi:10.1037/0882-7974.9.3.339

Lindquist, M. A., Meng Loh, J., Atlas, L. Y., & Wager, T. D. (2009). Modeling the hemodynamic response function in fMRI: Efficiency, bias and mis-modeling. *NeuroImage*, *45*(1), S187–S198. doi:10.1016/j.neuroimage.2008.10.065

Lippert, M., Logothetis, N. K., & Kayser, C. (2007). Improvement of visual contrast detection by a simultaneous sound. *Brain Research*, *1173*, 102–109. doi:10.1016/j.brainres.2007.07.050

Liu, H., Yang, Y., Xia, Y., Zhu, W., Leak, R. K., Wei, Z., … Hu, X. (2017). Aging of Cerebral White Matter. *Ageing Research Reviews*, *34*, 64–76. doi:10.1016/j.arr.2016.11.006

Liu, T. T., Glover, G. H., Mueller, B. A., Greve, D. N., & Brown, G. G. (2013). An Introduction to Normalization and Calibration Methods in Functional MRI. *Psychometrika*, *78*(2), 308–321. doi:10.1007/s11336-012-9309-x

Liu, X. Z., & Yan, D. (2007). Ageing and hearing loss. *The Journal of Pathology*, *211*(2), 188–197. doi:10.1002/path.2102

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878. doi:10.1038/nature06976

Lovelace, C. T., Stein, B. E., & Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Cognitive Brain Research*, *17*(2), 447–453. doi:10.1016/S0926-6410(03)00160-5

Luo, L., & Craik, F. I. (2008). Aging and Memory: A Cognitive Approach. *The Canadian Journal of Psychiatry*, *53*(6), 346–353. doi:10.1177/070674370805300603

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, *9*(11), 1432–1438. doi:10.1038/nn1790

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends in Neurosciences*, *28*(5), 264–271. doi:10.1016/j.tins.2005.03.008

Macaluso, E., Driver, J., & Frith, C. D. (2003). Multimodal Spatial Representations Engaged in Human Parietal Cortex during Both Saccadic and Manual Spatial Orienting. *Current Biology*, *13*(12), 990–999. doi:10.1016/S0960-9822(03)00377-4

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex in Cognitive Control. *Science*, *288*(5472), 1835–1838. doi:10.1126/science.288.5472.1835

MacDonald, J., Andersen, S., & Bachmann, T. (2000). Hearing by Eye: How Much Spatial Degradation can Be Tolerated? *Perception*, *29*(10), 1155–1168. doi:10.1068/p3020

Macleod, A., & Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *British Journal of Audiology*, *21*(2), 131–141. doi:10.3109/03005368709077786

Madden, D. J., & Allen, P. A. (1995). Aging and the Speed/Accuracy Relation in Visual Search: Evidence for an Accumulator Model. *Optometry and Vision Science*, *72*(3), 210.

Magnotti, J. F., & Beauchamp, M. S. (2017). A Causal Inference Model Explains Perception of the McGurk Effect and Other Incongruent Audiovisual Speech. *PLOS Computational Biology*, *13*(2), e1005229. doi:10.1371/journal.pcbi.1005229

Magnotti, J. F., Ma, W. J., & Beauchamp, M. S. (2013). Causal inference of asynchronous audiovisual speech. *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00798

Maguinness, C., Setti, A., Burke, K. E., Kenny, R. A., & Newell, F. N. (2011). The Effect of Combined Sensory and Semantic Components on Audio–Visual Speech Perception in Older Adults. *Frontiers in Aging Neuroscience*, *3*. doi:10.3389/fnagi.2011.00019

Mahoney, J. R., Dumas, K., & Holtzer, R. (2015). Visual–Somatosensory Integration is Linked to Physical Activity Level in Older Adults. *Multisensory Research*, *28*(1–2), 11–29. doi:10.1163/22134808-00002470

Mahoney, J. R., Holtzer, R., & Verghese, J. (2014). Visual-Somatosensory Integration and Balance: Evidence for Psychophysical Integrative Differences in Aging. *Multisensory Research*, *27*(1), 17–42. doi:10.1163/22134808-00002444

Mahoney, J. R., Li, P. C. C., Oh-Park, M., Verghese, J., & Holtzer, R. (2011). Multisensory integration across the senses in young and old adults. *Brain Research*, *1426*, 43–53. doi:10.1016/j.brainres.2011.09.017

Mahoney, J. R., Verghese, J., Dumas, K., Wang, C., & Holtzer, R. (2012). The effect of multisensory cues on attention in aging. *Brain Research*, *1472*, 63–73. doi:10.1016/j.brainres.2012.07.014

Mahoney, J. R., Verghese, J., Goldin, Y., Lipton, R., & Holtzer, R. (2010). Alerting, orienting, and executive attention in older adults. *Journal of the International Neuropsychological Society*, *16*(05), 877–889. doi:10.1017/S1355617710000767

Mahoney, J. R., Wang, C., Dumas, K., & Holtzer, R. (2014). Visual-somatosensory integration in aging: Does stimulus location really matter? *Visual Neuroscience*, *31*(03), 275–283. doi:10.1017/S0952523814000078

Maier, J. X., Di Luca, M., & Noppeney, U. (2011). Audiovisual asynchrony detection in human speech. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(1), 245–256. doi:10.1037/a0019952

Maiworm, M., & Roeder, B. (2011). Suboptimal auditory dominance in audiovisual integration of temporalcues. *Tsinghua Science and Technology*, *16*(2), 121–132. doi:10.1016/S1007-0214(11)70019-0

Maltby, M. T. (2002). *Principles of hearing aid audiology* (2nd ed). London: Whurr.

Mann, A. H., Graham, N., & Ashby, D. (1984). Psychiatric illness in residential homes for the elderly: A survey in one London borough. *Age and Ageing, 13*(5), 257–265. doi: 10.1093/ageing/13.5.257

Marks, B. L., Madden, D. J., Bucur, B., Provenzale, J. M., White, L. E., Cabeza, R., & Huettel, S. A. (2007). Role of Aerobic Fitness and Aging on Cerebral White Matter Integrity. *Annals of the New York Academy of Sciences*, *1097*(1), 171–174. doi:10.1196/annals.1379.022

Martuzzi, R., Murray, M. M., Michel, C. M., Thiran, J.-P., Maeder, P. P., Clarke, S., & Meuli, R. A. (2007). Multisensory Interactions within Human Primary Cortices Revealed by BOLD Dynamics. *Cerebral Cortex*, *17*(7), 1672–1679. doi:10.1093/cercor/bhl077

Matchin, W., Groulx, K., & Hickok, G. (2014). Audiovisual Speech Integration Does Not Rely on the Motor System: Evidence from Articulatory Suppression, the McGurk Effect, and fMRI. *Journal of Cognitive Neuroscience*, *26*(3), 606–620. doi:10.1162/jocn_a_00515

McGettigan, C., Faulkner, A., Altarelli, I., Obleser, J., Baverstock, H., & Scott, S. K. (2012). Speech comprehension aided by multiple modalities: Behavioural and neural interactions. *Neuropsychologia*, *50*(5), 762–776. doi:10.1016/j.neuropsychologia.2012.01.010

McGovern, D. P., Roudaia, E., Stapleton, J., McGinnity, T. M., & Newell, F. N. (2014). The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration. *Frontiers in Aging Neuroscience*, *6*. doi:10.3389/fnagi.2014.00250

McGovern, D., Roudaia, E., Stapleton, J., Li, X., Watson, D., McGinnity, T. M., & Newell, F. (2013). Age-related behavioural and neural differences in multisensory processing. In *Perception ECVP abstract* (Vol. 42, pp. 140–140). Retrieved from https://www.researchgate.net/profile/Tm_Mcginnity/publication/266157488_Age-related_behavioural_and_neural_differences_in_multisensory_processing/links/544eb7dd0cf26dda08901a24.pdf

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

Meijer, D., & Noppeney, U. (n.d.). Computational models of multisensory integration (p. 32).

Meltytech. (2018). *Shotcut*.

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, *214*(5–6), 655–667. doi:10.1007/s00429-010-0262-0

Meredith, M. A., & Stein, B. E. (1985). Spatial factors determine the activity of muitisensory neurons in cat superior colliculus. *Brain Research*, *365*, 5.

Meredith, M., & Stein, B. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, *221*(4608), 389–391. doi:10.1126/science.6867718

Merriman, N. A., Whyatt, C., Setti, A., Craig, C., & Newell, F. N. (2015). Successful balance training is associated with improved multisensory function in fall-prone older adults. *Computers in Human Behavior*, *45*, 192–203. doi:10.1016/j.chb.2014.12.017

Milham, M. P., Erickson, K. I., Banich, M. T., Kramer, A. F., Webb, A., Wszalek, T., & Cohen, N. J. (2002). Attentional Control in the Aging Brain: Insights from an fMRI Study of the Stroop Task. *Brain and Cognition*, *49*(3), 277–296. doi:10.1006/brcg.2001.1501

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*(2), 247–279. doi:10.1016/0010-0285(82)90010-X

Miller, L. M., & D'Esposito, M. (2005). Perceptual Fusion and Stimulus Coincidence in the Cross-Modal Integration of Speech. *Journal of Neuroscience*, *25*(25), 5884–5893. doi:10.1523/JNEUROSCI.0896-05.2005

Mohtasib, R. S., Lumley, G., Goodwin, J. A., Emsley, H. C. A., Sluming, V., & Parkes, L. M. (2012). Calibrated fMRI during a cognitive Stroop task reveals reduced metabolic response with increasing age. *NeuroImage*, *59*(2), 1143–1151. doi:10.1016/j.neuroimage.2011.07.092

Molholm, S., Sehatpour, P., Mehta, A. D., Shpaner, M., Gomez-Ramirez, M., Ortigue, S., … Foxe, J. J. (2006). Audio-Visual Multisensory Integration in Superior Parietal Lobule Revealed by Human Intracranial Recordings. *Journal of Neurophysiology*, *96*(2), 721–729. doi:10.1152/jn.00285.2006

Morcom, A. M., Henson, R. N. A., & Cam-CAN. (2018). Increased prefrontal activity with aging reflects nonspecific neural responses rather than compensation. *Journal of Neuroscience*, 1701–17. doi:10.1523/JNEUROSCI.1701-17.2018

Morcom, A. M., & Johnson, W. (2015). Neural Reorganization and Compensation in Aging. *Journal of Cognitive Neuroscience*, *27*(7), 1275–1285. doi:10.1162/jocn_a_00783

Morgan, M. L., DeAngelis, G. C., & Angelaki, D. E. (2008). Multisensory Integration in Macaque Visual Cortex Depends on Cue Reliability. *Neuron*, *59*(4), 662–673. doi:10.1016/j.neuron.2008.06.024

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human Primary Auditory Cortex: Cytoarchitectonic Subdivisions and Mapping into a Spatial Reference System. *NeuroImage*, *13*(4), 684–701. doi:10.1006/nimg.2000.0715

Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., & Laurienti, P. J. (2012). Multisensory integration and aging. In *The Neural Bases of Multisensory Processes*. Retrieved from http://europepmc.org/abstract/MED/22593868

Mullinger, K. J., Mayhew, S. D., Bagshaw, A. P., Bowtell, R., & Francis, S. T. (2014). Evidence that the negative BOLD response is neuronal in origin: A simultaneous EEG–BOLD–CBF study in humans. *NeuroImage*, *94*, 263–274. doi:10.1016/j.neuroimage.2014.02.029

Murray, D. J., Ellis, R. R., Bandomir, C. A., & Ross, H. E. (1999). Charpentier (1891) on the size—weight illusion. *Perception & Psychophysics*, *61*(8), 1681–1685. doi:10.3758/BF03213127

Nahorna, O., Berthommier, F., & Schwartz, J.-L. (2012). Binding and unbinding the auditory and visual streams in the McGurk effect. *The Journal of the Acoustical Society of America*, *132*(2), 1061–1077. doi:10.1121/1.4728187

Nahorna, O., Berthommier, F., & Schwartz, J.-L. (2015). Audio-visual speech scene analysis: Characterization of the dynamics of unbinding and rebinding the McGurk effect. *The Journal of the Acoustical Society of America*, *137*(1), 362–377. doi:10.1121/1.4904536

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. doi:10.1111/j.1532-5415.2005.53221.x

Nath, A. R., & Beauchamp, M. S. (2011). Dynamic Changes in Superior Temporal Sulcus Connectivity during Perception of Noisy Audiovisual Speech. *Journal of Neuroscience*, *31*(5), 1704–1714. doi:10.1523/JNEUROSCI.4853-10.2011

Nath, A. R., & Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *NeuroImage*, *59*(1), 781–787. doi:10.1016/j.neuroimage.2011.07.024

Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science*, *9*(5), 454–464. doi:10.1111/j.1467-7687.2006.00512.x

Newell, F. N., Bülthoff, H. H., & Ernst, M. O. (2003). Cross-modal perception of actively explored objects (p. 10).

Nikmat, A. W., Al-Mashoor, S. H., & Hashim, N. A. (2015). Quality of life in people with cognitive impairment: nursing homes versus home care. *International Psychogeriatrics, 27*(5), 815–824. doi:10.1017/S1041610214002609

Noesselt, T., Bergmann, D., Hake, M., Heinze, H.-J., & Fendrich, R. (2008). Sound increases the saliency of visual events. *Brain Research*, *1220*, 157–163. doi:10.1016/j.brainres.2007.12.060

Noppeney, U., Jones, S. A., Rohe, T., & Ferrari, A. (2018). See what you hear – How the brain forms representations across the senses. *Neuroforum.*

Noppeney, U., Ostwald, D., & Werner, S. (2010). Perceptual Decisions Formed by Accumulation of Audiovisual Evidence in Prefrontal Cortex. *Journal of Neuroscience*, *30*(21), 7434–7446. doi:10.1523/JNEUROSCI.0455-10.2010

Okada, K., Venezia, J. H., Matchin, W., Saberi, K., & Hickok, G. (2013). An fMRI Study of Audiovisual Speech Perception Reveals Multisensory Interactions in Auditory Cortex. *PLOS ONE*, *8*(6), e68959. doi:10.1371/journal.pone.0068959

Oruc, I., Sinnett, S., Bischof, W. F., Soto-Faraco, S., Lock, K., & Kingstone, A. (2008). The effect of attention on the illusory capture of motion in bimodal stimuli. *Brain Research*, *1242*, 200–208. doi:10.1016/j.brainres.2008.04.014

Ozker, M., Schepers, I. M., Magnotti, J. F., Yoshor, D., & Beauchamp, M. S. (2017). A Double Dissociation between Anterior and Posterior Superior Temporal Gyrus for Processing

Audiovisual Speech Demonstrated by Electrocorticography. *Journal of Cognitive Neuroscience*, *29*(6), 1044–1060. doi:10.1162/jocn_a_01110

Pagani, E., Agosta, F., Rocca, M. A., Caputo, D., & Filippi, M. (2008). Voxel-based analysis derived from fractional anisotropy images of white matter volume changes with aging. *NeuroImage*, *41*(3), 657–667. doi:10.1016/j.neuroimage.2008.03.021

Pandey, P. C., Kunov, H., & Abel, S. M. (1986). Disruptive effects of auditory signal delay on speech perception with lipreading. *Journal of Auditory Research*, *26*(1), 27–41.

Panek, P. E., Rush, M. C., & Slade, L. A. (1984). Locus of the Age-Stroop Interference Relationship. *The Journal of Genetic Psychology*, *145*(2), 209–216. doi:10.1080/00221325.1984.10532268

Parise, Cesare V., Spence, C., & Ernst, M. O. (2012). When Correlation Implies Causation in Multisensory Integration. *Current Biology*, *22*(1), 46–49. doi:10.1016/j.cub.2011.11.039

Parise, Cesare Valerio, & Spence, C. (2009). ‘When Birds of a Feather Flock Together’: Synesthetic Correspondences Modulate Audiovisual Integration in Non-Synesthetes. *PLOS ONE*, *4*(5), e5664. doi:10.1371/journal.pone.0005664

Park, D. C., & Reuter-Lorenz, P. (2009). The Adaptive Brain: Aging and Neurocognitive Scaffolding. *Annual Review of Psychology*, *60*(1), 173–196. doi:10.1146/annurev.psych.59.103006.093656

Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2018). Representational interactions during audiovisual speech entrainment: Redundancy in left posterior superior temporal gyrus and synergy in left motor cortex. *PLOS Biology*, *16*(8), e2006558. doi:10.1371/journal.pbio.2006558

Parker, J. L., & Robinson, C. W. (2018). Changes in multisensory integration across the life span. *Psychology and Aging*, *33*(3), 545–558. doi:10.1037/pag0000244

Patel, A. X., Kundu, P., Rubinov, M., Jones, P. S., Vértes, P. E., Ersche, K. D., … Bullmore, E. T. (2014). A wavelet method for modeling and despiking motion artifacts from resting-state fMRI time series. *NeuroImage*, *95*, 287–304. doi:10.1016/j.neuroimage.2014.03.012

Peelle, J. E. (2014). Methodological challenges and solutions in auditory functional magnetic resonance imaging. *Frontiers in Neuroscience*, *8*. doi:10.3389/fnins.2014.00253

Peelle, J. E., Troiani, V., Grossman, M., & Wingfield, A. (2011). Hearing Loss in Older Adults Affects Neural Systems Supporting Speech Comprehension. *Journal of Neuroscience*, *31*(35), 12638–12643. doi:10.1523/JNEUROSCI.2559-11.2011

Peiffer, A. M., Mozolic, J. L., Hugenschmidt, C. E., & Laurienti, P. J. (2007). Age-related multisensory enhancement in a simple audiovisual detection task. *Neuroreport*, *18*(10), 1077–1081.

Pekkola, J., Ojanen, V., Autti, T., Jääskeläinen, I. P., Möttönen, R., Tarkiainen, A., & Sams, M. (2005). Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *NeuroReport*, *16*(2), 125.

Persson, J., Nyberg, L., Lind, J., Larsson, A., Nilsson, L.-G., Ingvar, M., & Buckner, R. L. (2006). Structure–Function Correlates of Cognitive Decline in Aging. *Cerebral Cortex*, *16*(7), 907–915. doi:10.1093/cercor/bhj036

Peters, A., Morrison, J. H., Rosene, D. L., & Hyman, B. T. (1998). Feature article: are neurons lost from the primate cerebral cortex during normal aging? *Cerebral Cortex*, *8*(4), 295–300. doi:10.1093/cercor/8.4.295

Peters, Alan. (2002). The effects of normal aging on myelin and nerve fibers: a review. *Journal of Neurocytology*, *31*(8–9), 581–593.

Peters, M. A. K., Ma, W. J., & Shams, L. (2016). The Size-Weight Illusion is not anti-Bayesian after all: a unifying Bayesian account. *PeerJ*, *4*. doi:10.7717/peerj.2124

Pfefferbaum, A., Mathalon, D. H., Sullivan, E. V., Rawles, J. M., Zipursky, R. B., & Lim, K. O. (1994). A Quantitative Magnetic Resonance Imaging Study of Changes in Brain Morphology From Infancy to Late Adulthood. *Archives of Neurology*, *51*(9), 874–887. doi:10.1001/archneur.1994.00540210046012

Pitts, B. J., & Sarter, N. (2018). What You Don't Notice Can Harm You: Age-Related Differences in Detecting Concurrent Visual, Auditory, and Tactile Cues. *Human Factors*, 0018720818759102. doi:10.1177/0018720818759102

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of Functional MRI Data Analysis*. Cambridge University Press.

Poliakoff, E., Ashworth, S., Lowe, C., & Spence, C. (2006). Vision and touch in ageing: Crossmodal selective attention and visuotactile spatial interactions. *Neuropsychologia*, *44*(4), 507–517. doi:10.1016/j.neuropsychologia.2005.07.004

Presacco, A., Simon, J. Z., & Anderson, S. (2016). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *Journal of Neurophysiology*, *116*(5), 2346–2355. doi:10.1152/jn.00372.2016

Prins, N., & Kingdom, F. A. A. (2018). Applying the Model-Comparison Approach to Test Specific Research Hypotheses in Psychophysical Research Using the Palamedes Toolbox. *Frontiers in Psychology*, *9*. doi:10.3389/fpsyg.2018.01250

Rabbitt, P. (1979). How old and young subjects monitor and control responses for accuracy and speed. *British Journal of Psychology*, *70*(2), 305–311. doi:10.1111/j.2044-8295.1979.tb01687.x

Ratcliff, R. (2008). Modeling Aging Effects on Two-Choice Tasks: Response Signal and Response Time Data. *Psychology and Aging*, *23*(4), 900–916. doi:10.1037/a0013930

Raz, N., Gunning, F. M., Head, D., Dupuis, J. H., McQuain, J., Briggs, S. D., … Acker, J. D. (1997). Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. *Cerebral Cortex*, *7*(3), 268–282. doi:10.1093/cercor/7.3.268

Raz, N., Rodrigue, K. M., & Haacke, E. M. (2007). Brain Aging and Its Modifiers: Insights from in Vivo Neuromorphometry and Susceptibility Weighted Imaging. *Annals of the New York Academy of Sciences*, *1097*(1), 84–93. doi:10.1196/annals.1379.018

Raz, Naftali, Ghisletta, P., Rodrigue, K. M., Kennedy, K. M., & Lindenberger, U. (2010). Trajectories of brain aging in middle-aged and older adults: Regional and individual differences. *NeuroImage*, *51*(2), 501–511. doi:10.1016/j.neuroimage.2010.03.020

Raz, Naftali, Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., … Acker, J. D. (2005). Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers. *Cerebral Cortex*, *15*(11), 1676–1689. doi:10.1093/cercor/bhi044

Reisberg, D., McLean, J., & Goldfield, A. (1987). Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli. In *Hearing by eye: The psychology of lip-reading* (pp. 97–113). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Ren, Y. N., Yang, W., Takahashi, S., Nakahashi, K., & Wu, J. (2015). Temporal gap between visual and auditory stimuli lessen audiovisual integration in aging under cross-modal attention. In *2015 IEEE International Conference on Mechatronics and Automation (ICMA)* (pp. 511–516). doi:10.1109/ICMA.2015.7237538

Reuter-Lorenz, P. A., & Cappell, K. A. (2008). Neurocognitive Aging and the Compensation Hypothesis. *Current Directions in Psychological Science*, *17*(3), 177–182. doi:10.1111/j.1467-8721.2008.00570.x

Reuter-Lorenz, P. A., Marshuetz, C., Jonides, J., Smith, E. E., Hartley, A., & Koeppe, R. (2001). Neurocognitive ageing of storage and executive processes. *European Journal of Cognitive Psychology*, *13*(1–2), 257–278. doi:10.1080/09541440125972

Riecker, A., Grodd, W., Klose, U., Schulz, J. B., Gröschel, K., Erb, M., … Kastrup, A. (2003). Relation between Regional Functional MRI Activation and Vascular Reactivity to Carbon Dioxide during Normal Aging. *Journal of Cerebral Blood Flow & Metabolism*, *23*(5), 565–573. doi:10.1097/01.WCB.0000056063.25434.04

Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014). Bayesian model selection for group studies — Revisited. *NeuroImage*, *84*, 971–985. doi:10.1016/j.neuroimage.2013.08.065

Rohe, T., & Noppeney, U. (2015). Sensory reliability shapes perceptual inference via two mechanisms. *Journal of Vision*, *15*(5), 22.

Rohe, Tim, & Noppeney, U. (2015). Cortical Hierarchies Perform Bayesian Causal Inference in Multisensory Perception. *PLOS Biology*, *13*(2), e1002073. doi:10.1371/journal.pbio.1002073

Rohe, Tim, & Noppeney, U. (2016). Distinct Computational Principles Govern Multisensory Integration in Primary Sensory and Association Cortices. *Current Biology*, *26*(4), 509–514. doi:10.1016/j.cub.2015.12.056

Rohe, Tim, & Noppeney, U. (2018). Reliability-Weighted Integration of Audiovisual Signals Can Be Modulated by Top-down Control. *ENeuro*, ENEURO.0315-17.2018. doi:10.1523/ENEURO.0315-17.2018

Rosas, P., Wagemans, J., Ernst, M. O., & Wichmann, F. A. (2005). Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. *Journal of the Optical Society of America A*, *22*(5), 801. doi:10.1364/JOSAA.22.000801

Rosemann, S., & Thiel, C. M. (2018). Audio-visual speech processing in age-related hearing loss: Stronger integration and increased frontal lobe recruitment. *NeuroImage*. doi:10.1016/j.neuroimage.2018.04.023

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do You See What I Am Saying? Exploring Visual Enhancement of Speech Comprehension in Noisy Environments. *Cerebral Cortex*, *17*(5), 1147–1153. doi:10.1093/cercor/bhl024

Roudaia, E., Calabro, F. J., Vaina, L. M., & Newell, F. N. (2018). Aging Impairs Audiovisual Facilitation of Object Motion Within Self-Motion. *Multisensory Research*, *31*(3–4), 251–272. doi:10.1163/22134808-00002600

Roudaia, E., Sekuler, A. B., Bennett, P. J., & Sekuler, R. (2013). Aging and Audio-Visual and Multi-Cue Integration in Motion. *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00267

Sadaghiani, S., Maier, J. X., & Noppeney, U. (2009). Natural, Metaphoric, and Linguistic Auditory Direction Signals Have Distinct Influences on Visual Motion Processing. *Journal of Neuroscience*, *29*(20), 6490–6499. doi:10.1523/JNEUROSCI.5437-08.2009

Salat, D. H., Kaye, J. A., & Janowsky, J. S. (1999). Prefrontal Gray and White Matter Volumes in Healthy Aging and Alzheimer Disease. *Archives of Neurology*, *56*(3), 338. doi:10.1001/archneur.56.3.338

Salthouse, T. (1979). Adult age and the speed-accuracy trade-off. *Ergonomics*, *22*(7), 811–821. doi:10.1080/00140137908924659

Salthouse, T. (2012). Consequences of Age-Related Cognitive Declines. *Annual Review of Psychology*, *63*(1), 201–226. doi:10.1146/annurev-psych-120710-100328

Salthouse, T. A. (1991). Mediation of Adult Age Differences in Cognition by Reductions in Working Memory and Speed of Processing. *Psychological Science*, *2*(3), 179–183. doi:10.1111/j.1467-9280.1991.tb00127.x

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. doi:10.1037/0033-295X.103.3.403

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology*, *54*(1–3), 35–54. doi:10.1016/S0301-0511(00)00052-1

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society*, *16*(05), 754–760. doi:10.1017/S1355617710000706

Salthouse, T. A., Hancock, H. E., Meinz, E. J., & Hambrick, D. Z. (1996). Interrelations of Age, Visual Acuity, and Cognitive Functioning. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *51B*(6), P317-330. doi:10.1093/geronb/51B.6.P317

Sams, M., Möttönen, R., & Sihvonen, T. (2005). Seeing and hearing others and oneself talk. *Cognitive Brain Research*, *23*(2–3), 429–435. doi:10.1016/j.cogbrainres.2004.11.006

Schlack, A. (2005). Multisensory Space Representations in the Macaque Ventral Intraparietal Area. *Journal of Neuroscience*, *25*(18), 4616–4625. doi:10.1523/JNEUROSCI.0455-05.2005

Schmitz, T. W., Cheng, F. H. T., & Rosa, E. D. (2010). Failing to Ignore: Paradoxical Neural Effects of Perceptual Load on Early Attentional Selection in Normal Aging. *Journal of Neuroscience*, *30*(44), 14750–14758. doi:10.1523/JNEUROSCI.2687-10.2010

Schulte, T., Müller-Oehring, E. M., Chanraud, S., Rosenbloom, M. J., Pfefferbaum, A., & Sullivan, E. V. (2011). Age-related reorganization of functional networks for successful conflict resolution: A combined functional and structural MRI study. *Neurobiology of Aging*, *32*(11), 2075–2090. doi:10.1016/j.neurobiolaging.2009.12.002

Schütt, H. H., Harmeling, S., Macke, J. H., & Wichmann, F. A. (2016). Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Research*, *122*, 105–123. doi:10.1016/j.visres.2016.02.002

Schwartz, J.-L., Berthommier, F., & Savariaux, C. (2004). Seeing to hear better: evidence for early audio-visual interactions in speech identification. *Cognition*, *93*(2), B69–B78. doi:10.1016/j.cognition.2004.01.006

Sekiyama, K., Soshi, T., & Sakamoto, S. (2014). Enhanced audiovisual integration with aging in speech perception: a heightened McGurk effect in older adults. *Frontiers in Psychology*, *5*. doi:10.3389/fpsyg.2014.00323

Sekuler, Benett, & Mamelak. (2000). Effects of Aging on the Useful Field of View. *Experimental Aging Research*, *26*(2), 103–120. doi:10.1080/036107300243588

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*(6614), 308. doi:10.1038/385308a0

Setti, A., Burke, K. E., Kenny, R. A., & Newell, F. N. (2011). Is inefficient multisensory processing associated with falls in older people? *Experimental Brain Research*, *209*(3), 375–384. doi:10.1007/s00221-011-2560-z

Setti, A., Burke, K. E., Kenny, R., & Newell, F. N. (2013). Susceptibility to a multisensory speech illusion in older persons is driven by perceptual processes. *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00575

Setti, A., Finnigan, S., Sobolewski, R., McLaren, L., Robertson, I. H., Reilly, R. B., … Newell, F. N. (2011). Audiovisual temporal discrimination is less efficient with aging: an event-related potential study. *NeuroReport*, *22*(11), 554. doi:10.1097/WNR.0b013e328348c731

Setti, A., Stapleton, J., Leahy, D., Walsh, C., Kenny, R. A., & Newell, F. N. (2014). Improving the efficiency of multisensory integration in older adults: Audio-visual temporal discrimination training reduces susceptibility to the sound-induced flash illusion. *Neuropsychologia*, *61*, 259–268. doi:10.1016/j.neuropsychologia.2014.06.027

Shams, L., & Beierholm, U. R. (2010). Causal inference in perception. *Trends in Cognitive Sciences*, *14*(9), 425–432. doi:10.1016/j.tics.2010.07.001

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature*, *408*(6814), 788. doi:10.1038/35048669

Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport*, *16*(17), 1923–1927.

Shapleske, J., Rossell, S. ., Woodruff, P. W. ., & David, A. . (1999). The planum temporale: a systematic, quantitative review of its structural, functional and clinical significance. *Brain Research Reviews*, *29*(1), 26–49. doi:10.1016/S0165-0173(98)00047-2

Sharp, D. J., Scott, S. K., Mehta, M. A., & Wise, R. J. S. (2006). The Neural Correlates of Declining Performance with Age: Evidence for Age-Related Changes in Cognitive Control. *Cerebral Cortex*, *16*(12), 1739–1749. doi:10.1093/cercor/bhj109

Sinnett, S., Soto-Faraco, S., & Spence, C. (2008). The co-occurrence of multisensory competition and facilitation. *Acta Psychologica*, *128*(1), 153–161. doi:10.1016/j.actpsy.2007.12.002

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*, *17*(10), 2387–2399. doi:10.1093/cercor/bhl147

Sliwinski, M. J., & Hall, C. B. (1998). Constraints on general slowing: a meta-analysis using hierarchical linear models with random coefficients. *Psychology and Aging*, *13*(1), 164–175. doi:10.1037//0882-7974.13.1.164

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*, *12*(1), 7–10.

Smith, G. A., & Brewer, N. (1995). Slowness and age: Speed-accuracy mechanisms. *Psychology and Aging*, *10*(2), 238–247. doi:10.1037/0882-7974.10.2.238

Smith, J., & Baltes, P. B. (1990). Wisdom-related knowledge: Age/cohort differences in response to life-planning problems. *Developmental Psychology*, *26*(3), 494–505. doi:10.1037/0012-1649.26.3.494

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*(3), 199–222. doi:10.1023/B:STCO.0000035301.49549.88

Snell, K. B. (1997). Age-related changes in temporal gap detection. *The Journal of the Acoustical Society of America*, *101*(4), 2214–2220. doi:10.1121/1.418205

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging*, *14*(3), 458–472. doi:10.1037/0882-7974.14.3.458

Sommers, M. S., Tye-Murray, N., & Spehar, B. (2005). Auditory-Visual Speech Perception and Auditory-Visual Enhancement in Normal-Hearing Younger and Older Adults: *Ear and Hearing*, *26*(3), 263–275. doi:10.1097/00003446-200506000-00003

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience*, *6*(3), 309–315. doi:10.1038/nn1008

Spence, C. (2013a). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule: Spatial coincidence and multisensory integration. *Annals of the New York Academy of Sciences*, *1296*(1), 31–49. doi:10.1111/nyas.12121

Spence, C. (2013b). Multisensory flavour perception. *Current Biology*, *23*(9), R365–R369. doi:10.1016/j.cub.2013.01.028

Sperdin, H. F., Cappe, C., Foxe, J. J., & Murray, M. M. (2009). Early, low-level auditory-somatosensory multisensory interactions impact reaction time speed. *Frontiers in Integrative Neuroscience*, *3*. doi:10.3389/neuro.07.002.2009

Stapleton, J., Setti, A., Doheny, E. P., Kenny, R. A., & Newell, F. N. (2013). A standing posture is associated with increased susceptibility to the sound-induced flash illusion in fall-prone older adults. *Experimental Brain Research*, *232*(2), 423–434. doi:10.1007/s00221-013-3750-7

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging*, *25*(2), 377–390. doi:10.1037/a0018022

Staudinger, U. M., Smith, J., & Baltes, P. B. (1992). Wisdom-related knowledge in a life review task: Age differences and the role of professional specialization. *Psychology and Aging*, *7*(2), 271–281. doi:10.1037/0882-7974.7.2.271

Steffener, J., Brickman, A. M., Habeck, C. G., Salthouse, T. A., & Stern, Y. (2013). Cerebral blood flow and gray matter volume covariance patterns of cognition in aging. *Human Brain Mapping*, *34*(12), 3267–3279. doi:10.1002/hbm.22142

Stein, B. E., Huneycutt, W. S., & Meredith, M. A. (1988). Neurons and behavior: the same rules of multisensory integration apply. *Brain Research*, *448*(2), 355–358. doi:10.1016/0006-8993(88)91276-0

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*(4), 255–266. doi:10.1038/nrn2331

Stein, B. E., & Wallace, M. T. (1996). Chapter 20 Comparisons of cross-modality integration in midbrain and cortex. In *Progress in Brain Research* (Vol. 112, pp. 289–299). Elsevier. doi:10.1016/S0079-6123(08)63336-1

Stephen, J. M., Knoefel, J. E., Adair, J., Hart, B., & Aine, C. J. (2010). Aging-related changes in auditory and visual integration measured with MEG. *Neuroscience Letters*, *484*(1), 76–80. doi:10.1016/j.neulet.2010.08.023

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*(03), 448–460. doi:10.1017/S1355617702813248

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

Stevenson, R. A., Baum, S. H., Krueger, J., Newhouse, P. A., & Wallace, M. T. (2017). Links Between Temporal Acuity and Multisensory Integration Across Life Span. *Journal of Experimental Psychology. Human Perception and Performance*. doi:10.1037/xhp0000424

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), 1210–1223. doi:10.1016/j.neuroimage.2008.09.034

Stevenson, R. A., Nelms, C. E., Baum, S. H., Zurkovsky, L., Barense, M. D., Newhouse, P. A., & Wallace, M. T. (2015). Deficits in audiovisual speech perception in normal aging emerge at the level of whole-word recognition. *Neurobiology of Aging*, *36*(1), 283–291. doi:10.1016/j.neurobiolaging.2014.08.003

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), 1339–1345. doi:10.1016/j.neuroimage.2010.12.063

Streri, A., & Gentaz, E. (2003). Cross-modal recognition of shape from hand to eyes in human newborns. *Somatosensory & Motor Research*, *20*(1), 13–18. doi:10.1080/0899022031000083799

Sugihara, T., Diltz, M. D., Averbeck, B. B., & Romanski, L. M. (2006). Integration of Auditory and Visual Communication Information in the Primate Ventrolateral Prefrontal Cortex. *Journal of Neuroscience*, *26*(43), 11138–11147. doi:10.1523/JNEUROSCI.3550-06.2006

Sullivan, E. V., Rosenbloom, M., Serventi, K. L., & Pfefferbaum, A. (2004). Effects of age and sex on volumes of the thalamus, pons, and cortex. *Neurobiology of Aging*, *25*(2), 185–192. doi:10.1016/S0197-4580(03)00044-7

Sumby, W. H., & Pollack, I. (1954). Visual Contribution to Speech Intelligibility in Noise. *The Journal of the Acoustical Society of America*, *26*(2), 212–215. doi:10.1121/1.1907309

Swick, D., Ashley, V., & Turken, A. U. (2008). Left inferior frontal gyrus is critical for response inhibition. *BMC Neuroscience*, *9*(1), 102. doi:10.1186/1471-2202-9-102

Sutcliffe, C., Burns, A., Challis, D., Mozley, C. G., Cordingley, L., Bagley, H., & Huxley, P. (2007). Depressed Mood, Cognitive Impairment, and Survival in Older People Admitted to Care Homes in England. *The American Journal of Geriatric Psychiatry, 15*(8), 708–715. doi: 10.1097/JGP.0b013e3180381537

Szycik, G. R., Jansma, H., & Münte, T. F. (2009). Audiovisual integration during speech comprehension: An fMRI study comparing ROI-based and whole brain analyses. *Human Brain Mapping*, *30*(7), 1990–1999. doi:10.1002/hbm.20640

Takahashi, C., Diedrichsen, J., & Watt, S. J. (2009). Integration of vision and haptics during tool use. *Journal of Vision*, *9*(6), 3. doi:10.1167/9.6.3

Teasdale, N., & Simoneau, M. (2001). Attentional demands for postural control: the effects of aging and sensory reintegration. *Gait & Posture*, *14*(3), 203–210. doi:10.1016/S0966-6362(01)00134-5

Terry, R. D., DeTeresa, R., & Hansen, L. A. (1987). Neocortical cell counts in normal human adult aging. *Annals of Neurology*, *21*(6), 530–539. doi:10.1002/ana.410210603

Thomas, S. M., & Jordan, T. R. (2002). Determining the influence of Gaussian blurring on inversion effects with talking faces. *Perception & Psychophysics*, *64*(6), 932–944. doi:10.3758/BF03196797

Tiippana, K. (2014). What is the McGurk effect? *Frontiers in Psychology*, *5*. doi:10.3389/fpsyg.2014.00725

Tomer, R., & Levin, B. E. (1993). Differential Effects of Aging on Two Verbal Fluency Tasks. *Perceptual and Motor Skills*, *76*(2), 465–466. doi:10.2466/pms.1993.76.2.465

Townsend, J., Adamo, M., & Haist, F. (2006). Changing channels: An fMRI study of aging and cross-modal attention shifts. *NeuroImage*, *31*(4), 1682–1692. doi:10.1016/j.neuroimage.2006.01.045

Tremblay, C., Champoux, F., Voss, P., Bacon, B. A., Lepore, F., & Théoret, H. (2007). Speech and Non-Speech Audio-Visual Illusions: A Developmental Study. *PLoS ONE*, *2*(8), e742. doi:10.1371/journal.pone.0000742

Tsvetanov, K. A., Henson, R. N. A., Tyler, L. K., Davis, S. W., Shafto, M. A., Taylor, J. R., … Rowe, J. B. (2015). The effect of ageing on fMRI: Correction for the confounding effects of vascular reactivity evaluated by joint fMRI and MEG in 335 adults. *Human Brain Mapping*, *36*(6), 2248–2269. doi:10.1002/hbm.22768

Tsvetanov, K. A., Mevorach, C., Allen, H., & Humphreys, G. W. (2013). Age-related differences in selection by visual saliency. *Attention, Perception, & Psychophysics*, *75*(7), 1382–1394. doi:10.3758/s13414-013-0499-9

Turner, G. R., & Spreng, R. N. (2012). Executive functions and neurocognitive aging: dissociable patterns of brain activity. *Neurobiology of Aging*, *33*(4), 826.e1-826.e13. doi:10.1016/j.neurobiolaging.2011.06.005

Tye-Murray, N., Sommers, M., Spehar, B., Myerson, J., & Hale, S. (2010). Aging, Audiovisual Integration, and the Principle of Inverse Effectiveness. *Ear and Hearing*, *31*(5), 636–644. doi:10.1097/AUD.0b013e3181ddf7ff

Tye-Murray, N., Spehar, B., Myerson, J., Hale, S., & Sommers, M. (2016). Lipreading and Audiovisual Speech Recognition across the Adult Lifespan: Implications for Audiovisual Integration. *Psychology and Aging*, *31*(4), 380–389. doi:10.1037/pag0000094

Ulbrich, P., Churan, J., Fink, M., & Wittmann, M. (2009). Perception of Temporal Order: The Effects of Age, Sex, and Cognitive Factors. *Aging, Neuropsychology, and Cognition*, *16*(2), 183–202. doi:10.1080/13825580802411758

Vallet, G. T., Simard, M., Versace, R., & Mazza, S. (2013). The perceptual nature of audiovisual interactions for semantic knowledge in young and elderly adults. *Acta Psychologica*, *143*(3), 253–260. doi:10.1016/j.actpsy.2013.04.009

van Beers, R. J., Sittig, A. C., & van der Gon, J. J. D. (1999). Integration of Proprioceptive and Visual Position-Information: An Experimentally Supported Model. *Journal of Neurophysiology*, *81*(3), 1355–1364. doi:10.1152/jn.1999.81.3.1355

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2006). The Stroop Color-Word Test: Influence of Age, Sex, and Education; and Normative Data for a Large Sample Across the Adult Age Range. *Assessment*, *13*(1), 62–79. doi:10.1177/1073191105283427

van Doorn, J., Ly, A., Marsman, M., & Wagenmakers, E.-J. (2017). Bayesian Latent-Normal Inference for the Rank Sum Test, the Signed Rank Test, and Spearman's ρ. *ArXiv:1712.06941 [Stat]*. Retrieved from http://arxiv.org/abs/1712.06941

Van Gerven, P. W. M., & Guerreiro, M. J. S. (2016). Selective Attention and Sensory Modality in Aging: Curses and Blessings. *Frontiers in Human Neuroscience*, *10*. doi:10.3389/fnhum.2016.00147

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

Venezia, J. H., Vaden, K. I. J., Rong, F., Maddox, D., Saberi, K., & Hickok, G. (2017). Auditory, Visual and Audiovisual Speech Processing Streams in Superior Temporal Sulcus. *Frontiers in Human Neuroscience*, *11*. doi:10.3389/fnhum.2017.00174

Vercillo, T., & Gori, M. (2015). Attention to sound improves auditory reliability in audio-tactile spatial optimal integration. *Frontiers in Integrative Neuroscience*, *9*. doi:10.3389/fnint.2015.00034

Verhaeghen, P., & Cerella, J. (2008). Everything we know about aging and response times: A meta-analytic integration. In *Handbook of cognitive aging: Interdisciplinary perspectives* (pp. 134–150). Thousand Oaks, CA, US: Sage Publications, Inc. doi:10.4135/9781412976589.n8

Vroomen, J., Bertelson, P., & De Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & Psychophysics*, *63*(4), 651–659. doi:10.3758/BF03194427

Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1583–1590. doi:10.1037//0096-1523.26.5.1583

Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition*, *40*(2), 145–160. doi:10.3758/s13421-011-0158-0

Walhovd, K. B., Fjell, A. M., Reinvang, I., Lundervold, A., Dale, A. M., Eilertsen, D. E., … Fischl, B. (2005). Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiology of Aging*, *26*(9), 1261–1270. doi:10.1016/j.neurobiolaging.2005.05.020

Wallace, M. T., Meredith, M. A., & Stein, B. E. (1993). Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *Journal of Neurophysiology*, *69*(6), 1797–1809. doi:10.1152/jn.1993.69.6.1797

Wallace, M. T., Wilkinson, L. K., & Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *Journal of Neurophysiology*, *76*(2), 1246–1266. doi:10.1152/jn.1996.76.2.1246

Walton, J. P., Frisina, R. D., & O'Neill, W. E. (1998). Age-Related Alteration in Processing of Temporal Sound Features in the Auditory Midbrain of the CBA Mouse. *Journal of Neuroscience*, *18*(7), 2764–2776. doi:10.1523/JNEUROSCI.18-07-02764.1998

Wang, L., Mruczek, R. E. B., Arcaro, M. J., & Kastner, S. (2015). Probabilistic Maps of Visual Topography in Human Cortex. *Cerebral Cortex*, *25*(10), 3911–3931. doi:10.1093/cercor/bhu277

Ward, L. M., Aitchison, R. T., Tawse, M., Simmers, A. J., & Shahani, U. (2015). Reduced Haemodynamic Response in the Ageing Visual Cortex Measured by Absolute fNIRS. *PLOS ONE*, *10*(4), e0125012. doi:10.1371/journal.pone.0125012

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*(8), 989–994. doi:10.1016/S0028-3932(02)00316-0

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*(8), 1293–1313. doi:10.3758/BF03194544

Wiley, T. L., Chappell, R., Carmichael, L., Nondahl, D. M., & Cruickshanks, K. J. (2008). Changes in Hearing Thresholds over 10 Years in Older Adults. *Journal of the American Academy of Audiology*, *19*(4), 281–371.

Wilson, R. S., Bienias, J. L., Evans, D. A., & Bennett, D. A. (2004). Religious Orders Study: Overview and Change in Cognitive and Motor Speed. *Aging, Neuropsychology, and Cognition*, *11*(2–3), 280–303. doi:10.1080/13825580490511125

Winneke, A. H., & Phillips, N. A. (2011). Does audiovisual speech offer a fountain of youth for old ears? An event-related brain potential study of age differences in audiovisual speech perception. *Psychology and Aging*, *26*(2), 427–438. doi:10.1037/a0021683

Wong, P. C. M., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., & Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia*, *47*(3), 693–703. doi:10.1016/j.neuropsychologia.2008.11.032

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*(1), 58–73. doi:10.1002/(SICI)1097-0193(1996)4:1<58::AID-HBM4>3.0.CO;2-O

Wozny, D. R., Beierholm, U. R., & Shams, L. (2010). Probability Matching as a Computational Strategy Used in Perception. *PLoS Computational Biology*, *6*(8), e1000871. doi:10.1371/journal.pcbi.1000871

Wright, L. A., & Wormald, R. P. L. (1992). Stereopsis and ageing. *Eye*, *6*(5), 473–476. doi:10.1038/eye.1992.100

Wu, J., Yang, W., Gao, Y., & Kimura, T. (2012). Age-related multisensory integration elicited by peripherally presented audiovisual stimuli. [Miscellaneous Article]. *Neuroreport July 11, 2012*, *23*(10), 616–620. doi:10.1097/WNR.0b013e3283552b0f

Yu, A. J., Dayan, P., & Cohen, J. D. (2009). Dynamics of Attentional Selection Under Conflict: Toward a Rational Bayesian Account. *Journal of Experimental Psychology*, *35*(3), 700–717. doi:10.1037/a0013553

Zacks, R. T., Hasher, L., & Li, K. Z. H. (2000). Human memory. In *The handbook of aging and cognition, 2nd ed* (pp. 293–357). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Zanto, T. P., & Gazzaley, A. H. (2014). Attention and Ageing. In *The Oxford Handbook of Attention*.

Zhu, D. C., Zacks, R. T., & Slade, J. M. (2010). Brain activation during interference resolution in young and older adults: An fMRI study. *NeuroImage*, *50*(2), 810–817. doi:10.1016/j.neuroimage.2009.12.087

Zimpfer, V., & Sarafian, D. (2014). Impact of hearing protection devices on sound localization performance. *Frontiers in Neuroscience*, *8*. doi:10.3389/fnins.2014.00135

Zou, Z., Chau, B. K. H., Ting, K.-H., & Chan, C. C. H. (2017). Aging Effect on Audiovisual Integrative Processing in Spatial Discrimination Task. *Frontiers in Aging Neuroscience*, *9*. doi:10.3389/fnagi.2017.00374

Zweig, L. J., Suzuki, S., & Grabowecky, M. (2014). Learned face–voice pairings facilitate visual search. *Psychonomic Bulletin & Review*, *22*(2), 429–436. doi:10.3758/s13423-014-0685-3

Zysset, S., Schroeter, M. L., Neumann, J., & Yves von Cramon, D. (2007). Stroop interference, hemodynamic response and aging: An event-related fMRI study. *Neurobiology of Aging*, *28*(6), 937–946. doi:10.1016/j.neurobiolaging.2006.05.008