LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH
Ollscoil Átha Cliath

TRINITY COLLEGE LIBRARY DUBLIN
The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Elucidating Multisensory Interactions in Synaesthesia

by

Gary Bargary

A dissertation submitted for the degree of Doctor of Philosophy of the University of Dublin, Trinity College, Dublin 2, Ireland.

This research was conducted in the School of Psychology and in the Institute of Neuroscience.

June 2008

Declaration

I hereby declare that this thesis:

a) has not been submitted as an exercise for a degree at this or any other University,

b) comprises the results of my own investigations.

I give permission to the Library to lend or copy this thesis upon request, subject to normal conditions of acknowledgement.

Signed:

Gary Bargary

Date: 18/06/08

Summary

This work examines synaesthesia; a neurodevelopmental phenomenon widely believed to be the result of aberrant cross-talk from one cortical area to another. A number of behavioural measures were used to assess the nature of the aberrant cross-talk in synaesthesia. Also, 'normal' cross-talk between the senses was examined to gain insight into the mechanisms which enable cross-talk or multisensory interactions between the senses.

The first issue addressed the question of whether the cross-talk in linguistic-colour synaesthesia is specific to their synaesthesia or if indeed there are broader cross-talk differences in synaesthetes. As a measure of cross-talk between the senses a number of audiovisual integration experiments were carried out by the synaesthetes. The findings showed that the cross-talk is not completely specific to their synaesthesia, but instead complex audiovisual stimuli (such as speech) showed greater cross-talk (basic stimuli such as beeps and flashes did not show any difference). The argument was made that this may be due to 'special' status of linguistic entities in synaesthesia.

The next issue addressed the question of where in the hierarchy of processing does the synaesthetic cross-talk manifests. The findings from this study indicated that the synaesthetic experience was not tied to low-level properties of the inducing stimuli (i.e. stimuli which in induce synaesthetic experiences). This suggests that the cross-talk may be at a higher cognitive level, rather than being between two low-level sensory cortices. The last issue was concerned with how 'normal' cross-talk occurs, examining a form of pseudo-synaesthesia present in everyone. The findings from this study suggested that the brain does not care about the information carrier (i.e. the sense or modality) but rather the content of the information, and consequently an amodal approach (information is extracted directly across multiple sources) to cross-talk and multisensory interactions was taken.

This amodal perspective and the other findings obtained led to a re-evaluation of cross-talk theories which posit map to map (sensory or conceptual) cross-activation. Instead the argument was made that the development of linguistic skills may play a critical role in synaesthesia. More specifically, linguistic-colour synaesthesia may be the outcome of a deviation in the higher-level capacity of forming perceptual anchors. Such an account provides a reason of why linguistic entities are so dominant in synaesthesia, a necessary requirement for any theory of synaesthesia.

Acknowledgements

I would first like to thank my supervisors, Fiona Newell and Kevin Mitchell for invaluable advice and help throughout this project. The past and present research team, Jason, Kylie, Andy, Sarah, Aisling, Joanne, Iwona and Christina also provided lots of bits and pieces of advice and help along the way, particularly Jason and Kylie who I owe a great deal. I would also like to thank David Hevey, Hugh Garavan and my external examiner Jamie Ward for many helpful comments.

My twin brother Brian I also owe a great deal and without his advice and support this thesis would not have been possible. And finally I owe the greatest debt to my girlfriend Aileen for supporting me (and putting up with me) while carrying out and writing up this work.

Publications Arising from the Present Work

Bargary, G., Mitchell, K. J., Barnett, K. J. & Newell, F. (submitted). The McGurk illusion provides evidence for a late, perceptually-driven onset of synaesthesia.

Bargary, G., Chan, J. & Newell, F. (submitted). Seeing Where the Ears Hear: Visual Encoding of Auditory Spatial Sequences.

Bargary, G. & Mitchell, K. J. (in press). Synaesthesia and cortical connectivity. *Trends in Neuroscience*.

Table of Contents

Declaration....................i

Summary....................ii

Acknowledgements....................iv

Publications Arising from the Present Work....................v

Table of Contents....................vi

List of Figures....................xi

List of Tables....................xiii

Chapter 1. Introduction....................1

1.1 Why Study Synaesthesia?....................2

1.2 Types, Trends and Variability in Synaesthesia....................5

1.3 Familiality, Prevalence and Genetics of Synaesthesia....................6

1.4 Behavioural Studies....................8

1.4.1 Validating Synaesthesia....................8

1.4.2 At What Stage in Processing Does Synaesthesia Occur?....................11

1.4.2.1 Bottom Up or Top Down Influences?....................12

1.4.2.2 Is Attention Required?....................13

1.5 Imaging of Neural Activation in Synaesthetes.....................15

1.6 Neural Models....................17

1.7 Unanswered Questions....................21

1.8 Direction of Thesis....................22

Chapter 2. Investigating Audiovisual Integration in Synaesthetes....................25

2.1 Experiment 1....................30

2.1.1 Methods....................33

2.1.1.1 Participants....................33

2.1.1.2 Apparatus and Stimuli....33

2.1.1.3 Design. 34

2.1.1.4 Procedure. 35

2.1.2 Results and Discussion 36

2.2 Experiment 2....39

2.2.1 Methods....40

2.2.1.1 Participants....40

2.2.1.2 Stimuli....41

2.2.1.3 Design. 41

2.2.1.4 Procedure. 42

2.2.2 Results and Discussion 42

2.3 Experiment 3....44

2.3.1 Methods....45

2.3.1.1 Participants....45

2.3.1.2 Stimuli and apparatus....45

2.3.1.3 Design. 48

2.3.1.4 Procedure. 48

2.3.2 Results and Discussion 49

2.4 General discussion 53

Chapter 3. Synaesthesia and McGurk Effect 59

3.1 Experiment 4....60

3.1.1 Methods....61

3.1.1.1 Participants....61

3.1.1.2 Stimuli and apparatus....61

3.1.1.3 Design. 62

3.1.1.4 Procedure. 62

3.1.2 Results...62

3.1.3 Discussion...67

Chapter 4. The Role of Grapheme Processing in Linguistic-Colour Synaesthesia 73

4.1 Experiment 5(a) 77

4.1.1 Methods...78

4.1.1.1 Participants...78

4.1.1.2 Stimuli...78

4.1.1.3 Design. 79

4.1.1.4 Procedure. 80

4.1.2 Results and Discussion 80

4.2 Experiment 5(b) 84

4.2.1 Methods...84

4.2.1.1 Participants...84

4.2.1.2 Stimuli...85

4.2.1.3 Design. 85

4.2.1.4 Procedure. 85

4.2.2 Results and Discussion 86

4.3 Experiment 6(a) 88

4.3.1 Methods...90

4.3.1.1 Participants...90

4.3.1.2 Stimuli and apparatus..90

4.3.1.3 Design. 91

4.3.1.4 Procedure. 92

4.3.2 Results and Discussion 92

4.4 Experiment 6(b) 95
  4.4.1 Methods..96
    4.4.1.1 Participants...96
    4.4.1.2 Stimuli..96
    4.4.1.3 Design. 97
    4.4.1.4 Procedure. 97
  4.4.2 Results and Discussion 97
4.5 General Discussion 100
Chapter 5. Examining 'Normal' Cross-Talk Between the Senses. 104
5.1 Experiment 7 108
  5.1.1 Methods..110
    5.1.1.1 Participants...110
    5.1.1.2 Apparatus and stimuli. 110
    5.1.1.3 Design. 113
    5.1.1.4 Procedure. 114
  5.1.2 Results and Discussion 114
5.2 Experiment 8 116
  5.2.1 Methods..117
    5.2.1.1 Participants...117
    5.2.1.2 Apparatus and stimuli. 117
    5.2.1.3 Design. 117
    5.2.1.4 Procedure. 117
  5.2.2 Results and Discussion 119
5.3 Experiment 9 120
  5.3.1 Methods..121

5.3.1.1 Participants......121

5.3.1.2 Apparatus and stimuli......121

5.3.1.3 Design......121

5.3.1.4 Procedure......121

5.3.2 Results and Discussion......122

5.4 General discussion......124

Chapter 6. Discussion and Conclusions......129

6.1 Summary of the Results......129

6.2 Discussion of Overall Results......130

6.3 Implications for Neural Models......134

6.4 Is language Special in Synaesthesia?......137

6.5 Perceptual Anchors......138

6.6 Future Directions......141

6.7 Synopsis......142

References......144

List of Figures

Figure 1: Neural Connectivity Models of Synaesthesia. 18

Figure 2: The apparatus and stimuli set-up used in Experiment 1..............................34

Figure 3. Plot illustrating percentage of illusions for both flash and beep illusions (fission and fusion illusions) across synaesthetes and controls (Experiment 1)..........38

Figure 4: Plot illustrating unimodal variances (sd^2) for the all the beep and flash responses across synaesthetes and controls (Experiment 1). 38

Figure 5: Plot illustrating percentage of illusory responses reported across different stimulus onset asynchronies between two beeps, for both synaesthetes and controls (Experiment 2). 42

Figure 6: The visual and audio component of each audiovisual speech stimulus for each condition (Experiment 3 & 4)...48

Figure 7: Plot illustrating proportion of visual or fusion response across each condition (Experiment 3). 50

Figure 8: Plot illustrating proportion of correct audio response across each condition (Experiment 3). 53

Figure 9: Plot illustrating RGB colour difference for three different response set comparisons (Experiment 4). 64

Figure 10: Plot illustrating RGB colour difference for two different response set comparisons (Experiment 4). 65

Figure 11: Plot illustrating RGB colour difference, per individual synaesthete, in RGB vector distance using the same comparison as in Figure 10 (Experiment 4)..............66

Figure 12: Plot illustrating reaction time data across each of the three SOAs (stimulus onset asynchronies) and the conditions (Experiment 5(a))..81

Figure 13: Plot illustrating proportion correct across each of the two SOA's (stimulus onset asynchronies) and each condition (Experiment 5(a)). 83

Figure 14: Plot illustrating reaction time data across each of the two SOA's (stimulus onset asynchronies) and the four conditions (Experiment 5(b)). 86

Figure 15: Plot illustrating proportion correct across each of the two SOA's (stimulus onset asynchronies) and the four conditions (Experiment 5(b)). 88

Figure 16: Plot illustrating mean colour naming times across each SOA for the 'no interference' condition (Experiment 6(a)). 93

Figure 17: Plot illustrating the mean colour naming times across both SOAs and both interference conditions (Experiment 6(a)). 95

Figure 18: Plot illustrating mean colour naming times across both congruent and incongruent trials for each interference stimulus (Experiment 6(b)). 99

Figure 19: The apparatus used in all three experiments (Experiment 7-9). 111

Figure 20: The temporal profile of the auditory and visual stimuli in Experiment 7. 113

Figure 21: Plot illustrating the mean percentage correct performance in an audio-visual matching task across temporal disparities of 250 and 300 ms (Experiment 7). 115

Figure 22: Plot illustrating the mean percentage correct performance across the four different conditions (Experiment 8). 119

Figure 23: Plot illustrating the mean percentage correct performance in across the three conditions (Experiment 9). 122

Figure 24: Plot illustrating the mean percentage correct performance in Experiment 9 plotted in chronological order across bins of 5 trials. 124

List of Tables

Table 1: List of experimental word stimuli and corresponding expected percepts used in Experiment 3....47

Chapter 1. Introduction

Synaesthesia is a familial condition in which particular sensory stimuli can elicit the experience of a colour, taste or some other basic percept in the absence of environmental stimulation that would normally evoke such percepts (Rich & Mattingley, 2002). For example, the sound of a word may elicit (as well as the normal percept accompanying a spoken word) an additional percept of a taste, smell or colour, depending on the type of synaesthesia one has.

This condition has been known for the last 200 years (Cytowic, 2002), and was a hot topic in psychology at around the turn of the 20th century but interest declined in the rest of this century. The subjective nature of synaesthesia and the strict behaviourist climate which arose this century most likely contributed to its decline. Often psychologists and neuroscientists alike would dismiss the condition, attributing it too childhood associations, metaphorically speak, or the use of drugs (LSD and mescaline have synaesthetic-like effects) (Shannon, 2002). However, with the improvement of experimental techniques and imaging technology, synaesthesia research has undergone something of a renaissance in recent years (Ward & Mattingley, 2006).

There exist many different types of synaesthesia, which vary according to the range of stimuli which can induce a synaesthetic experience (called the inducer) and also the type of anomalous experience induced (also known as the concurrent) (Grossenbacher & Lovelace, 2001). Common inducers include musical sounds (Ward, Tsakanikos & Bray, 2006), linguistic elements (letters, words, numbers, time units) (Rich, Bradshaw & Mattingley, 2005), while rarer inducers include general sounds, odours, pain, personalities or temperatures. Among the types of concurrent, colour and specific spatial arrangements are the most common, while less common varieties include taste, smell, and touch (Day, 2005).

Given these differences there are however many similarities between synaesthetes unusual perceptual association which aid in characterising the condition; the associations are specific, present from an early age, stable, unidirectional (at least consciously, (Cohen Kadosh *et al.*, 2005), and idiosyncratic to each synaesthete. For example a grapheme-colour synaesthete may have a red experience upon presentation of the letter 'A', 'A' will always have this specific colour as far back as the synaesthete can remember, the colour red does not evoke the perception of the letter 'A' and the particular association will be unique to that synaesthete – synaesthetes rarely agree on their individual associations (however, a number of regularities in the associations have been observed (Barnett *et al.*, 2007; Rich *et al.*, 2005; Simner *et al.*, 2005). Also, it must be added that most synaesthetes view their synaesthesia as a welcome enrichment in their lives and would not be free of it given the choice.

1.1 Why Study Synaesthesia?

Synaesthesia is a very unique and curious phenomenon. A condition where certain people see the world differently than most is intriguing, and for some, is reason enough to study such a condition, but is its study worthy of scientific resources? Many people report quirky phenomenon such as usual ways of remembering numbers or doing math problems, is reporting certain colours for letters similar, is it little more than a "benign cognitive variant" (see Ward & Mattingley, 2006). One reason to reject this idea is that synaesthesia is very common; it is estimated that its prevalence could be as high as 4% (Simner, Mulvenna *et al.*, 2006). If it is simply a cognitive variant one has to explain why it keeps arising. Similarly, the familial nature of synaesthesia and its unique characteristics – it’s developmental, specific, stable, and unidirectional – suggest that this phenomenon is far from a simple cognitive variant but that it is a robust condition which can develop in the brain of certain individuals in its unique form, consistently.

Also, synaesthesia given it unique characteristic can help inform theories of ‘normal’ cognition or brain development (Ward & Mattingley, 2006). One topic which synaesthesia may shed light on is the process of multisensory integration or how the senses talk to each other. Similarities between normal cross-modal associations and synaesthetic associations have been observed (Grossenbacher & Lovelace, 2001; Marks, 1975; Sagiv & Ward, 2006; Smilek, Carriere, Dixon & Merikle, 2007; Ward, Huckstep, & Tsakanikos, 2006) and have led certain researchers to argue that synaesthesia is a case of heightened integration (Ward *et al.*, 2006). A related topic sometimes alluded to as synonymous with the previous one, is that synaesthesia may tell us something about the problem of binding (Mulvenna & Walsh, 2006; Robertson, 2003; Sagiv & Robertson, 2005; Weiss, Zilles & Fink. 2005), this is the idea the information from the different modalities needs to be combined or bound together in the brain to give us a unified perceptual experience, synaesthesia then is seen as a case of hyper-binding (Esterman, Verstynen, Ivry & Robertson, 2006; Hubbard, 2007). Studying synaesthesia for these researchers may then shed light on how contents of consciousness are bound together and thus how consciousness arises in the brain. However, whether ‘binding’ is indeed a problem the brain has to solve or whether it is in fact a pseudo-problem (O’Regan & Noë, 2001), i.e. a problem which evaporates under a different analysis or approach to consciousness, is much debated (Dennett, 1991; Noë, 2005).

Recently, for related reasons a group of philosophers have taken interest in synaesthesia, addressing issues of consciousness in cases of intra-modal and cross-modal plasticity, such as the late Susan Hurley and Alva Noë (2003). They found synaesthesia to be a difficult case for their perceptual theory (Hurley and Noë, 2003; Noë & Hurley, 2003; but c.f. Hurley & Noë, 2006). The late Jeffrey Gray (2003) bore out this difficulty and argued that synaesthesia cannot be explained by their perceptual theory and similar philosophical theories (Gray, 2003; Gray *et al.*, 2006).

Outside of issues of consciousness and multisensory perception, synaesthesia can inform theories of neurodevelopment, more specifically cortical development and plasticity. It is widely agreed that synaesthesia is the result of abnormal cross-talk from one cortical area of the brain to another (Hubbard & Ramachandran, 2005; Ward & Mattingley, 2006), how this situation in synaesthesia is set up in the brain can tell us more about certain neurodevelopment mechanisms and how they operate. For example, it is known that there is strong influence of the environment in shaping and refining cortical areas (Krubitzer & Kahn, 2003) and yet synaesthetic concurrents remain throughout life. Why don't the normal plasticity mechanisms driven by environmental input allow adaptation and the disappearance of these non-veridical experiences in synaesthesia? Also, whatever the developmental mechanisms are that lead to synaesthesia, how can they act so specifically to affect apparently one small portion of the brain? Why don't synaesthetes show more anomalies i.e. a broader phenotype? Related to neurodevelopment and plasticity are cases of acquired synaesthesia, where individuals report synaesthetic like effects after drug use or prolonged sensory deprivation (often caused by peripheral neural damage). For example, Ramachandran and Hubbard (2001, pg. 11) cite an example of a patient with retinitis pigmentosa who became progressively blind from childhood and by the age of 40 was completely blind. Interestingly a few years later he started experiencing tactile sensations as simple visual sensations, tactile input here seems to be, as Ramachandran and Hubbard suggest, invading and activating visual areas, see also Jacobs, 1981; Ro *et al.*, 2007; Ward, 2007). Understanding how such phenomena (both acquired and developmental synaesthesia) arise and how they are maintained gain insight into plasticity mechanisms in the brain.

Also, synaesthesia can be used to study aspects of linguistic processes (Simner, 2007), numerical cognition (Cohen Kadosh & Henik, 2007) or indeed space and magnitude perception (Hubbard, Piazza, Pinel & Dehaene, 2005). Thus, synaesthesia research can be informative for many diverse topics i.e. perception, cognition, cortical development and consciousness. Indeed, a thorough understanding of synaesthesia may very well will require an advancements in our understanding of the processes in each of these domains.

1.2 Types, Trends and Variability in Synaesthesia

The different types of synaesthesia vary depending on the inducer-concurrent pair. Numerous forms of synaesthesia have been reported, a study conducted by Sean Day (2005) surveyed a total of 572 synaesthetes and observed 35 different types of synaesthesia based on inducer-concurrent pairings. Inducers varied from graphemes, time units, musical sounds, general sounds, musical notes, phonemes, tastes, odours, pain, personalities, touch, temperatures and orgasms, whereas concurrents varied colour, smell, sound, taste, visual forms, and personality. Quantifying how many types of synaesthesia there are is difficult as it depends on how you characterize the inducing stimulus. For example, a broad characterization of the inducer such as linguistic-colour synaesthesia, could be given a finer grain characterisation on closer inspection where the inducer could be the grapheme, phoneme, morpheme, the word, lexical stress or lexical semantics (Simner, 2007). However, what is clear is that there is a wide degree of variability in the types of synaesthesia one can acquire.

Even within a certain type of synaesthesia there can be considerable variability in the nature of the concurrent induced. Synaesthetes can report their synaesthetic concurrent as existing 'out in the world' or in there 'minds eye', and are grouped as projectors and associators respectively (Dixon, Smilek & Merikle, 2004). Also, evidence exists that these different sub-types can be distinguished by their performance on behavioural tasks (Dixon *et al*, 2004; Hubbard, Arman, Ramachandran, & Boynton, 2005; Ward, Li, Salih, & Sagiv, 2006). Interestingly, within these many different types and possible sub-types of synaesthesia, there exist specific trends in the distribution of various types of synaesthesia observed. Linguistic inducers are very common, reported as high as approximately 88% of all synaesthesias (Simner, 2007), while colour and spatial arrangements are very common concurrents (Simner, Mulvenna *et al.*, 2006). Pain, smell or sounds are rarely concurrents (Day, 2005; Barnett el al., 2007). Why these trends exist is unknown, but it suggests that certain types of synaesthesia are favoured somehow either due to genetic, developmental or possibly cognitive reasons.

1.3 Familiality, Prevalence and Genetics of Synaesthesia

As alluded to in the introduction synaesthesia is known to be a familial trait, recent studies looking at grapheme-colour synaesthesia by Rich *et al.* (2005) and Ward and Simner (2005) showed a high incidence of synaesthesia within families, 36% and 44% respectively. A number of studies have shown also that synaesthesia is more common in females than in males, with both a UK and Australian based study reporting 6:1 ratio's (Baron-Cohen, Burt, Smith-Laittan, Harrison, & Bolton, 1996; Rich *et al.*, 2005). However recent work by Simner, Mulvenna *et al.* (2006) showed that when a large scale random sample study was conducted in a science museum the female bias disappeared, suggesting that the initial female bias reported in the literature may be confounded by a self report bias i.e. females are more likely to respond to health surveys (Dindia & Allen, 1992). Consistent with this hypothesis was a study carried out by Ward and Simner (2005) where in a self-referred sample of synaesthetes the female bias was 4:1 but when the synaesthetes relatives were surveyed for synaesthesia the female bias dropped to 2:1. A more recent larger study, however by Barnett *et al.*, (2007) showed a 6:1 ratio in the synaesthetes relatives where no self report bias was present, more research is needed to determine what the exact ratio is.

Another contention point in synaesthesia research is the prevalence of the condition. There have been many estimates ranging from 1 in 20 (Galton, 1883) to 1 in 25,000 (Cytowic, 1989), but the most cited estimate is 1/2000 (Baron-Cohen *et al.*, 1996). Recently a study carried out by Simner, Mulvenna *et al.*, (2006) which did not rely on self referral as many of the previous estimates did, suggests that synaesthesia could be as common as 1/20 for all forms of synaesthesia and 1/100 for grapheme-colour synaesthesia.

The mode of inheritance in synaesthesia is also not determined; it was originally suggested to be an x-linked dominant trait with a possible 50% lethality rate in utero for males (Baron-Cohen *et al.*, 1996) which could explain the high female bias and the lack of father-son transmission which is expected from an x-linked trait. More recent studies have also found no case of father-son transmission (Cytowic, 2002; Rich *et al.*, 2005; Ward & Simner, 2005) consistent with an X-linked gene.

However, studies looking at the number of siblings born out of synaesthetic mothers have found no evidence for male lethality (Barnett *et al*, 2007; Ward and Simner, 2005). Also preliminary results of genetic analysis do not implicate the X-chromosome in synaesthesia (Asher, personal communication). Even without male lethality in utero, it is still possible that the gene/genes may be x-linked given the uncertainty over the sex-ratio, but clearly further research is needed. Recently, Eagleman and others (2007) collecting data from over a 1000 synaesthetes have carried out a linkage analysis and have identified a region on chromosome 16q12.2-23.1, which may be the locus of trait.

1.4 Behavioural Studies

Many of the original behavioural studies on synaesthesia were concerned with verifying the existence of the phenomenon. To that end, many studies were concerned with showing that synaesthesia was automatic and indeed sensory or perceptual in nature. Further efforts investigating the nature of synaesthesia were very much concerned with what stage of processing synaesthesia manifested.

1.4.1 Validating Synaesthesia

The first measure of synaesthesia, known as the 'Test of Genuineness' (Baron-Cohen, Wyke & Binnie, 1987), established that the synaesthetes were genuine and weren't simply making up their synaesthetic associations. It involved measuring the consistency of a synaesthetes reported associations over a test-retest period of a month or a year. Synaesthetes typically get scores of 90-100 % consistency, while controls perform at 10-20% consistency (when asked to make arbitrary associations). This level of consistency is much greater than what would be expected from memory alone is what would be expected if these associations are in fact genuine and not confabulations.

A more recent behavioural measure which clearly distinguishes synaesthetes from non-synaesthetes and verifies the automatic nature of synaesthesia, uses a modified version of the Stroop task (Woolen & Ruggiero, 1983; Dixon, Smilek, Cudahy, & Merikle, 2000). In the standard Stroop task colour names are presented in an ink colour either congruent to that colour name or incongruent to the colour name. When participants are asked to name the colour of the ink reactions times are faster for the congruent trials as opposed to incongruent trials. This occurs due to the automatic nature of reading, as even though they are not instructed to read the word they read it automatically and this interferes with there response as regards naming the ink colour (Macleod, 1991). In the 'synaesthetic stroop' task graphemes are either presented in an ink colour congruent or incongruent to their synaesthetic colours. Synaesthetes tend to be quicker to name the ink colour of the graphemes in the congruent condition than in the incongruent condition. Another variation of the 'synaesthetic stroop' task uses instead of a coloured grapheme a colour patch, which is either congruent or incongruent to an inducer presented just prior. Again, interference from the synaesthetic colour is observed and the synaesthetes are quicker to name the colour of the patch in the congruent condition. This effect has been observed in wide variety of experimental tasks with synaesthetes and has verified the automatic nature of synaesthesia (Beeli, Essien, & Jancke, 2005; Dixon *et al*., 2000; Mattingley, Rich, Yelland & Bradshaw, 2001; Mills, Boteler & Oliver, 1999; O'dgaard, Flowers & Bradman, 1999; Wollen & Ruggiero, 1983).

The next primary concern was to show that the synaesthetic experiences were perceptual in nature rather than learned associations. As 'the synaesthetic Stroop' alone does demonstrate the automaticity of synaesthesia it does not address the nature of the synaesthetic associations, as it has been shown that stroop interference can occur with learned associations (Elias, Saucier, Hardie, & Sarty, 2003; MacLead & Dunbar, 1988). The general approach by researchers was to try and show that synaesthetic colours behaved like real colours in perceptual tasks. Ramachandran and Hubbard (2001b) provided evidence that synaesthetic colours behaved in a similar (but not identical) manner to normal colours in perceptual grouping and texture segregation tasks for two grapheme-colour synaesthetes. For example, in their texture segregation task, visual displays were briefly presented in which one of four shapes (square, triangle, rectangle, diamond) composed of target graphemes were embedded among distracter graphemes (the target and the distracter graphemes elicited different synaesthetic colours). The task was to indicate which of the four shapes was present by pressing the corresponding button. The two synaesthetes were significantly more accurate than controls. This study was further expanded to include six synaesthetes and the results showed that five out of the six synaesthetes were significantly more accurate than controls at identifying the shapes (Hubbard, Arman *et al.*, 2005). These improved performances in perceptual tasks suggest that for the synaesthetes studied that the synaesthetic colours induced improved performance similarly to how real colours would have, thus compatible with the idea that the synaesthetic experiences are of a perceptual nature.

Further studies on single grapheme-colour synaesthetes corroborated these findings using the similar approach of trying to show that synaesthetic colour behaved like real colours. Smilek and others (Smilek, Dixon, Cudahy, & Merikle, 2001; Smilek, Dixon & Merikle, 2003) reported two experiments where a synaesthete was slower to identify and localise a grapheme when its synaesthetic colour was congruent with the background colour than when it was incongruent. Palmeri and others (Palmeri, Blake, Marois & Whetsell, 2002) also showed by modifying a standard visual search task, that visual search times for identifying a target grapheme measured against number of distracters (16 or 25 or 36 distracters) was more efficient (less of an increase in reaction time as distracter number increases) when the target and distracter graphemes were a different colour than to when they were the same synaesthetic colour. More recent studies by the same group also showed that synaesthetic colours can behave like real colour in having various grouping effects and the waterfall illusion (Kim & Blake, 2005; Kim, Blake & Palmeri, 2006). Even though these earlier studies clearly showed a difference in perceptual tasks between synaesthetes and non-synaesthetes and went some way to validating the perceptual reality of synaesthetic colours, more recent studies involving visual search tasks (Edquist, Rich, Brinkman & Mattingley, 2005: Laeng, Svatdal & Oelmann, 2004) have cast doubt on the exact interpretation of these earlier results (see section 1.4.2.2).

1.4.2 At What Stage in Processing Does Synaesthesia Occur?

A fundamental question in the next step to understand what is different about synaesthetes and what is the exact nature of the synaesthetic associations, was to find out at what stage of processing synaesthesia occurs. For example in grapheme-colour synaesthesia, is lower-level sensory information of a grapheme enough to initiate a synaesthetic experience or is the synaesthetic experience initiated further downstream in the processing hierarchy? Possibly at an abstract level of the grapheme, or some other higher-order process involved in language processing. Related to this question, is focused attention required to induce the synaesthesia or can a synaesthetic concurrent be initiated without attending to the synaesthetic inducer. It has been taken that if synaesthesia can occur pre-attentively this provides evidence for synaesthesia occurring at a early stage of sensory processing. The answers to these questions are essential for developing neural or cognitive models of synaesthesia.

1.4.2.1 Bottom Up or Top Down Influences? Evidence for the role of top-down influences in initiating a synaesthetic experience comes from a variety of sources. Dixon and colleagues (Dixon *et al.*, 2000) showed that conceptual information alone without physical presentation of the inducing stimulus was enough to induce a synaesthetic experience. The study involved presenting an arithmetic math problem in sequence, where in place of the answer of the math problem a colour patch congruent or incongruent with the answer was presented. Synaesthetes were quicker at identifying the colour of the patch when the colour was congruent to the synaesthetic colour of the answer, demonstrating that the concept of the number was enough to initiate the synaesthetic colour. Further studies demonstrated this phenomenon conclusively (Dixon, Smilek, Duffy, Zanna & Merikle, 2006; Jansari, Spiller & Redfern, 2006). Another study, which similarly highlighted the importance of conceptual information or meaning in synaesthesia showed that the synaesthetic colour of an ambiguous grapheme i.e. a grapheme that could be interpreted as a number or a letter, depended on the context the grapheme was presented (Myles *et al.*, 2003).

Recent work by Simner, Glover & Mowat (2006) has shown that both bottom up and top-down influences can exist in determining the synaesthetic colour of a word. The synaesthetic colour of a word is often determined by the salient graphemes in the word (Baron-Cohen *et al.*, 1996), either the first letter or a dominant vowel in the word. Simner *et al.*, showed that the stress of the word was critical (e.g. 'con-vict versus con-'vict) and this was independent of whether the inducer was spoken or written words, indicating a possible effect of conceptual information. In another experiment they argued for an effect of bottom-up processes, as they found that synaesthetes were quicker to name the synaesthetic colour of the word ‘ether’ rather than ‘ethos’, indicating that the different graphemes compete for dominance in determining the synaesthetic colour of a word i.e. the e’s in ether facilitated the synaesthetic colour of the word quicker than in ethos where there is a competing e and o (see also Simner, 2007).

Further evidence for the role of bottom-up processes in inducing synaesthetic concurrents comes from evidence that specific features of the inducer such as contrast (Hubbard, Manohar & Ramachandran, 2006) can effect the reported strength of the synaesthetic colours induced. Also Witthoft and Winawer (2006) showed that the saturation of synaesthetic colours depended on the font and case of letter stimuli for one synaesthete.

1.4.2.2 Is Attention Required? Determining whether synaesthesia requires focused attention or whether it can be induced pre-attentively has been one of the main controversies in synaesthesia research. The work described above by Ramchandran and Hubbard (2001b), Palmeri *et al*., (2002) and Smilek *et al*., (2001, 2003) which showed that synaesthetic colours can act somewhat like real colours in perceptual tasks, suggested that synaesthetic exhibit the same ‘pop-out’ effect as real colours, indicative of pre-attentive automatic processing. Particularly in the Palmeri *et al*., (2002) study the search slopes for the synaesthete was almost parallel indicating a pre-attentive parallel search.

However, there is ample evidence for an involvement of attention in synaesthesia. Mattingley *et al*., (2001) showed in a study involving 15 grapheme-colour synaesthetes that by eliminating the conscious awareness of the grapheme using a masking paradigm, this eliminated the synaesthetic colour as measured by the 'synaesthetic stroop' effect. Similarly in another study by Mattingley *et al.*, (2003), which used navon-type displays where a grapheme is composed entirely of much smaller graphemes, showed that whether the synaesthetes (14 synaesthetes were tested) attended locally or globally influenced the synaesthetic colour perceived as measured by the 'synaesthetic stroop' effect.

Laeng *et al.*, (2004) carried out an interesting single case study on a grapheme-colour synaesthete using a visual search paradigm, which demonstrated, in agreement with Palmeri *et al.*, (2002) and a pre-attentive interpretation, a more efficient search (less of an increase in reaction times as distracter size increased) than controls when target and distracter graphemes elicited different colours. Interestingly, however they also found evidence for a role of attention in their study. A closer look at the data revealed that synaesthetes only performed in a pre-attentive manner when the target grapheme was near the point of fixation (at eccentricities of 3° and 6° visual angle), but when the target was at greater eccentricities no pre-attentive affect is found. They concluded that visual search was only facilitated when the grapheme was in the spotlight of attention, so it was not a pre-attentive pop-out search, but is instead attention mediated. They draw the same conclusion as regards Palmeri and others (2002) results. A recent study on 14 synaesthetes by Edquist *et al.* (2006) using a similar paradigm as the Palmeri *et al.*, (2002) study (except that this study used a between-subjects design rather than a within subject design) found no pre-attentive affects (i.e. identical performance regardless of the amount of distracters).

It has been suggested that the heterogeneity of synaesthesia (Dixon *et al.*, 2004; Dixon & Smilek, 2005; Hubbard, Arman *et al.*, 2005;) and the different tasks employed in each study, may explain the contradictory results obtained from these studies (Edquist *et al.*, 2006; Hubbard & Ramachandran, 2005).

1.5 Imaging of Neural Activation in Synaesthetes.

Many of the early neuroimaging studies were concerned with validating the existence of synaesthesia, particularly as, if the colour area hV4 (Zeki & Marini, 1998) was found to be active in linguistic-colour synaesthetes when they were presented with inducers; this would provide strong evidence for the existence of synaesthetic concurrents. The first imaging study was carried out by Paulesu *et al.*, (1995) using PET (Positron Emission Topography). Linguistic-colour synaesthetes were presented with either pure tones or single words, when synaesthetes heard words as opposed to tones areas of parieto-occipital junction and the posterior-inferior temporal cortex (PIT) were more active than in controls (today PIT may also be referred to as the visual word form area (Cohen & Dehaene, 2004)). Interestingly however, early visual areas V1, V2 and hV4 were not found to be significantly active upon listening to words compared with tones. Two single case fMRI (functional Magnetic Resonance Imaging) studies carried out in 2001 found early visual activation. Weiss and others (Weiss, Shah, Toni, Zilles & Fink 2001) found increased activation in extra striate cortex (near hV4), in a synaesthete who reported seeing colours upon hearing names, and Aleman and others (Aleman, Rutten, Sitskoorn, Dautzenberg & Ramsey, 2001) found activation in V1 but the authors were not able to determine if hV4 was active.

The first study to find hV4 activation in linguistic-colour synaesthetes was Nunn *et al* (2002). They tested six female linguistic-colour synaesthetes, who listened to words and tones. The colour area hV4 was found more active in the synaesthetes when they listened to words rather than tones, than in 6 matched controls. Even when the controls were trained up to imagine certain colours for certain words hV4 was not active. However, in another more recent fMRI study Weiss *et al*., 2005, enhanced hV4 activation was not observed in synaesthetes, instead regions of the intraparietal sulcus were differentially active. The authors concluded that this area could be essential for the anomalous 'binding' of synaesthetic inducer and concurrent. Hubbard *et al*., (2005) carried out an fMRI study with 6 grapheme-colour synaesthetes and 6 matched controls which found hV4 activation similar to the Nunn *et al*., study. Interestingly, the degree of activation in hV4 and other retinotopic areas (V1, V2 and V3) correlated with performance on perceptual tasks, subjects with better performance showed greater activation in these areas (however this greater activation did not reach significance in V1 and V2). Another recent study (Sperling, Prvulovic, Linden, Singer & Stirn, 2006) using a different methodology – it used graphemes that elicited no colour i.e. the synaesthetic colour was white, black or grey for the particular letter presented and compared them with graphemes that did elicit colour. They found increased activation in hV4 in the synaesthetes when they observed graphemes that elicited colours. The recent fMRI imaging study (Rich *et al*., 2006), however did not find any differential activation in hV4 when synaesthetes viewed graphemes but instead found activation in left medial lingual gyrus which is implicated in colour knowledge (Rich *et al*,. 2006).

It hard to reconcile these conflicting results, even though hV4 activation has been observed in a number of studies an equal number of studies don't find any hV4 activation and implicate other regions which could be necessary for synaesthesia – retinotopic visual areas (Hubbard, Arman *et al*., 2005), intraparietal sulcus (Weiss *et al*., 2005) or left medial lingual gyrus (Rich *et al*., 2006), each of which argues why each of these regions may be crucial for synaesthesia. As in the behavioural studies, individual differences and methodological differences may account for some of the inconsistencies.

1.6 Neural Models

All contemporary models of synaesthesia agree that synaesthesia is the result of increased cross-talk form one cortical area to another, either through structural or functional cortical connectivity differences (but c.f. Cytowic, 2002). The main point of contention is how exactly the concurrent perception is stimulated or induced (for a recent review see Hubbard and Ramachandran, 2005).

Two routes have been proposed (See Figure 1), a direct route – where the inducer directly cross-activates the concurrent, and an indirect route where the inducer activates the concurrent indirectly through another higher-order cortical area due to feedback connections. The nature of this anomalous activation is also subject to controversy, whether it is extra axonal connections (structural) or whether it is due to disinhibition of already present circuitry (functional). Typically a structural difference is coupled with the direct route i.e. Ramachandran's and Hubbard's (2001) cross-activation model, and a functional difference is coupled with the indirect model i.e. the disinhibition model (Grossenbacher & Lovelace, 2001; Ward *et al.*, 2006) or the re-entrant model (Dixon *et al.*, 2000). However there is no logical reason why this needs to be the case, the indirect model is equally compatible with a structural difference (as suggested by Ward *et al.*, 2006) whereas a direct model is compatible with a disinhibition imbalance. See Figure 1 for an illustration of these alternate neural connectivity models.

Evidence for an indirect versus direct route comes from studies showing that semantic information may be necessary for a synaesthetic experience (Dixon *et al.*, 2000; Dixon *et al.*, 2006; Myles *et al.*, 2003) and thus feedback from areas such as posterior intraparietal cortex may activate the concurrent.

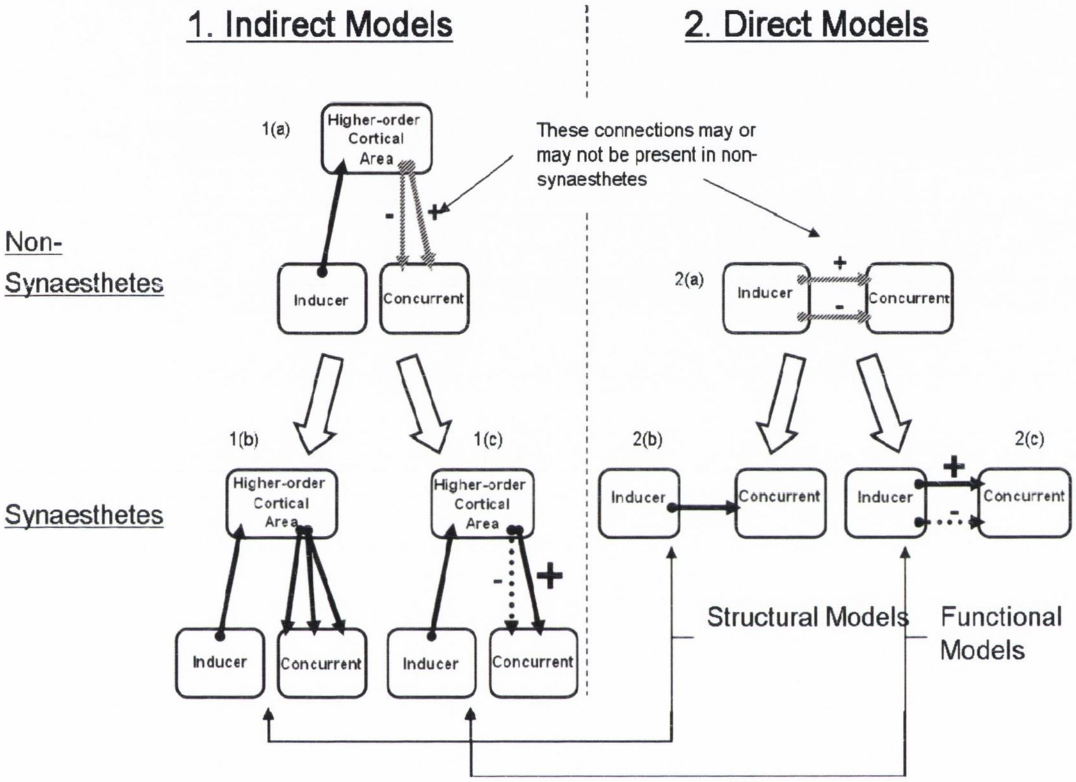

Figure 1: Neural Connectivity Models of Synaesthesia.
1. Cases where feedback from higher-order cortices are implicated, 1(a) non-synaesthetes may or may not have these feedback connections. In synaesthetes this feedback can be axonal connections (1(b)) or disinhibition (1 (c)). 2. Cases where there is direct cross-activation from one adjacent cortical area to another, 2(a) this cross-activation may or may not be present in non-synaesthetes. In synaesthetes the cross-activation can be either structural (axonal) (2(b)) or functional (2 (c)).

Also, evidence that the intraparietal sulcus may play a role in synaesthesia has come from two recent repetitive Transcranial Magnetic Stimulation (rTMS) studies. They found that when rTMS was applied to the right intraparietal sulcus of each synaesthete that this disrupted their synaesthesia as measured by the absence of the synaesthetic stroop effect (Esterman, Verstynen, Ivry & Robertson, 2006; Muggleton, Tsakanikos, Walsh. & Ward, 2007). This result supports previous suggestions that feedback from areas involved in integrating information activate the concurrent (Grossenbacher & Lovelace, 2001; Ward *et al*., 2006; Weiss *et al*., 2005).

Whereas the direct route model has to it advantage its simplicity and an existing synaesthetic like phenomenon known to be caused by adjacent invasion of axons, i.e. in some arm amputees the face area of the cortex invades the neighbouring deafferentiated arm area in somatosensory cortex, with the result that touching particular parts of the face induce particular feelings across the phantom limb (Ramachandran & Hirstein, 1998). The most substantial evidence to date however for structural differences in the brains of synaesthetes is a recent DTI study showing increased connectivity (inferred from measured fractional anisotropic) in right fusiform gyrus, left intraparietal sulcus and frontal areas of synaesthetes (Rouw & Scholte, 2007). Support for a functional difference (without any structural differences) in synaesthetes comes from cases of acquired synaesthesia. As mentioned previously, after peripheral nerve damage (Jacobs *et al*., 1981) producing synaesthetic-like phenomenon. It is argued that this is due to unmasking of existing connections where synaesthesia 'proper' could arise from a similar 'functional unmasking'. However, it is also possible given the time course of the phenomenon that the synaesthetic-like effects could be due to sprouting of new connections or even both functional and structural changes (Hubbard & Ramachandran, 2005). Further evidence for a functional difference comes from synaesthetic-like phenomena which have been reported when individuals are under the influence of psychedelic drugs (Shanon, 2002), given the timescale involved this would indicate functional rather than structural changes. Also it has been shown that certain synaesthetic cross-modal mechanisms are similar to non-synaesthetic cross-modal behaviour, leading researchers to argue that synaesthetes simply express a more extreme version of these normal cross-modal mechanisms (Marks, 1975; Ward *et al.*, 06). This view it is suggested is more compatible with a functional difference in synaesthetes rather than a structural difference. In the main the functional models claim that the connectivity that causes synaesthesia is present in us all but simply masked by a correct balance of inhibition and excitation, synaesthesia arises when this balance is upset.

Recently, another model has been proposed, which tries to accommodate both the recent DTI evidence (Rouw & Scholte, 2007) for structural differences in synaesthetes and the recent TMS studies (Esterman *et al.*, 2006; Muggleton *et al.*, 2007). It involves a two-stage process to explain linguistic-colour synaesthesia, the first stage involves direct cross-activation of fusiform areas as in the cross-activation model of Ramachandran and Hubbard (2001b), whereas the second step involves a 'binding' of the individual features of the inducer-concurrent pair in the parietal cortex (Hubbard, 2007).

Given the heterogeneity which is often reported in synaesthesia (Dixon *et al.*, 2004; Dixon & Smilek, 2005; Hubbard, Arman *et al.*, 2005; Marks, 1975), searching for a 'one model fits all' could indeed be a futile attempt (Hubbard & Ramachandran, 2005; Rich & Mattingley, 2002;), however given the evidence that a single genetic mechanism underlies many different types of synaesthesia, this makes it quite likely that the same underlying developmental mechanism exists but manifests itself differently among synaesthetes. For example, it is unlikely that the same developmental mechanism could develop as a functional difference for one synaesthete and a structural difference for another, as the genes involved in say axonal guidance are more than likely different from the genes involved in the determining the correct balance between inhibition and excitation.

1.7 Unanswered Questions

The most pertinent question at present, in trying to understand how synaesthesia arises in the brain, is to identify the gene/genes involved in synaesthesia. Identifying the nature of the gene/genes i.e. the protein that gene encodes, can tell us whether it is involved in axonal guidance, involved in the balance between excitation and inhibition or some other as yet unexplored (in regards to synaesthesia research) neural mechanisms i.e. a gene involved in plasticity for example. Once the gene is identified this will also increase our understanding of whatever neurodevelopmental is implicated, either by cloning the gene and looking at its effect on the mouse or by simply showing us what develops when the gene/genes is mutated. Also identifying a gene or cluster of genes unique to all varieties of synaesthesia would confirm that synaesthesia is not cluster of related conditions which involve different genes and different neurodevelopmental mechanisms (Marks & Odgaard, 2005; Rich and Mattingley, 2002), which would agree with data from pedigree analysis where different types of synaesthesia can run in the same family, thus suggesting a single genetic mechanism (Barnett *et al.*, 2007).

Another unanswered question is that given that we have so many types of synaesthesia what makes some types more common than others, particular why are explicitly learned entities such as letters and numbers so commonly inducers and why is colour so commonly a concurrent? Is there something special about linguistic phenomenon that makes them more likely to become inducers? Also, why aren't other areas of the brain affected by the genetic factor(s) behind synaesthesia, why don't we see increased cross-talk in other areas of the brain i.e. why don't we see a broader or endo-phenotype (Gottesman & Gould, 2003) in synaesthesia. It has to be determined also if the areas involved in synaesthesia are next to each other in the brain i.e. if adjacency is a defining characteristic of synaesthesia. Adjacency is compatible with a number of forms of synaesthesia (Hubbard & Ramachandran, 2007), so much so that the idea of adjacent cross-activation has been suggested to be able to account for all types of synaesthesia – the grand unified theory of synaesthesia (Hubbard, Simner & Ward, 2007).

The question of where in the hierarchy of processing synaesthesia occurs, or where the cross-talk manifests itself, is still unclear. Most of the evidence does favour that a certain amount of attentional resources are required, suggesting a higher-level manifestation of the synaesthetic association, but whether synaesthesia can be completely sensory-driven or whether a more higher-level (possibly linguistic) information is involved has yet to be resolved.

1.8 Direction of Thesis

This thesis will examine various aspects of synaesthesia and normal cross-talk between the senses using a number of behavioural measures. The first question addressed (Chapter 2), was do synaesthetes show increased (or decreased) cross-talk between the senses in general. As a measure of cross-talk between the senses various multisensory tasks were employed looking at the degree of multisensory integration in each. Multisensory integration is when two or more senses combine or are unified to give a unique experience different from each modality alone. The ventriloquist illusion is popular example of audiovisual integration. The synaesthetes were compared to controls in their performance on three audiovisual integration tasks, examining various aspects of multisensory integration. This experiment addresses two issues: whether synaesthetes exhibit more cross-talk between the senses in general due to knock on effects of the synaesthetic gene i.e. do synaesthetes show a broader phenotype and whether synaesthetes may use common cross-modal mechanisms but simply an extreme version of normal cross-modal mechanisms. If this is the case synaesthetes may exhibit more cross-modal integration than non-synaesthetes.

The next study (Chapter 3) on synaesthetes looked at what stage in the hierarchy of processing synaesthesia arises. More specifically it asked the question of whether synaesthesia is driven by early sensory input or late perceptual output. A well known multisensory illusion called the 'McGurk' effect (McGurk & MacDonald, 1976) was employed to answer this question. The final study on synaesthetes (Chapter 4) examined whether synaesthetes when presented with auditory inducers (i.e. speech) underwent a second step of grapheme conversion of this speech stimuli in order for the synaesthetic colour to be activated. Given the requirement of adjacent cortical areas in a recent model of synaesthesia it was argued that for auditory inducers to induce a synaesthetic colour a secondary step requiring grapheme conversion was required (Hubbard *et al.*, 2007). This hypothesis was tested directly using an interference paradigm.

The final experimental chapter (Chapter 5) examined a sort of pseudo-synaesthesia which was discovered recently by Guttman, Gilroy, and Blake (2005). They showed that a gabor patch alternating contrasts at a specific rate induces automatic auditory impressions – subjects report 'hearing' the rhythm of the visual stimulus. Guttman and others (2005) referred to such a phenomenon as a case of pseudo-synaesthesia in that visual temporal information is inducing 'hearing' of visual temporal rhythms'. They argued that because audition is better at dealing with temporal information that vision is, that it took over the task and thus vision inforamtion was encoded into an auditory format. I carried a similar investigation but focused instead on spatial information and how vision could influence auditory spatial encoding. This task enabled an investigation of the normal processes at work in the communication between the senses and also an examination of the properties which facilitate this cross-modal transfer of information. Ultimately this kind of investigation may shed light on the aspects of any inducer (both in synaesthesia proper and in this kind of pseudo-synaesthesia) which enable it (or make it more likely) to set up cross-modal associations.

Chapter 2. Investigating Audiovisual Integration in Synaesthetes

Synaesthesia is a curious neurodevelopmental condition for many reasons, chief among these, but yet an under-explored one, is the fact that synaesthetes do not appear to show any other significant behavioural or cognitive anomalies (Ward & Mattingley, 2005). The genetic and developmental factors that give rise to synaesthesia appear to act specifically, presumably only affecting certain cortical areas. This is complicated by the fact that amongst this developmental specificity there is extensive variability in the type of synaesthesia one can acquire. How, if the causal factors behind synaesthesia act specifically, can certain individuals develop one type of synaesthesia such as linguistic colour synaesthesia (Simner *et al.*, 2006) while others may develop linguistic-gustatory synaesthesia (Ward & Simner, 2003), both of which would involve diverse cortical areas.

One explanation for this variability is that synaesthesia itself may not represent a single neurodevelopmental condition but is instead the manifestation of a plethora of related conditions each involving different genes and possibly different neurodevelopmental mechanisms (Marks & Odgaard, 2005; Rich & Mattingley, 2002). An individual synaesthete can have multiple forms of synaesthesia (Cytowic, 1989) suggesting instead that each type of synaesthesia may be linked at a developmental level. It is a possibility that synaesthetes with multiple types may simply have inherited multiple genetic variants each effecting different developmental mechanisms. However, evidence from pedigree analysis (Barnett *et al.*, 2007) showing that different types and sub-types (Dixon *et al.*, 2004) of synaesthesia can run in the same family, provide strong evidence that a single genetic and developmental mechanism underlie all types of synaesthesia. Anecdotal reports of different types of synaesthesia running in the same family corroborate pedigree analysis (Blakemore, Bristow, Bird, Frith & Ward, 2005; Ward, Simner & Auyeung, 2005).

A singular genetic and developmental mechanism suggests that what synaesthetes actually inherit is a predisposition to develop synaesthesia, where epigenetic, developmental noise (Mitchell, 2007) or environmental factors determine which type of synaesthesia develops (Barnett *et al.*, 2007). It also suggests that early in development the synaesthetic brain may contain, or at least has the potential to contain, widespread connectivity differences (either structural or functional in nature) and that certain factors acting selectively (in terms of cortical location) during development either correct or prevent these connectivity differences from becoming established in the mature brain.

It is also a possibility, however that maybe widespread connectivity differences *do* exist in synaesthetes and that the synaesthetic condition is not as specific as the standard phenotypic analysis implies. This situation could arise from the genetic factor(s) (a single developmental mechanism does not imply a single gene) behind synaesthesia having knock-on effects elsewhere in the brain or it could come out of interactions early in development between non-differentiated functional modules (Huttenlocher & de Courten, 1987; Huttenlocher & Dabholkar, 1997) but which later in development become separate modules or processing systems (Johnson, 2001). Indeed, these latter developmental interactions could have a number of side-effects, which from the vantage point of examining the end-state of the condition would not be readily apparent (Thomas & Karmiloff-Smith, 2002). Only an analysis which takes the developmental trajectory into account could make sense of such extra anomalies (Karmiloff-Smith, 1998; Scerif & Karmiloff-Smith, 2005). Indeed in taking such a perspective (Ansari & Karmiloff-Smith, 2002; Karmiloff-Smith, 1998) on development (i.e. paying attention to the process of development itself) one would expect that research would uncover not only anomalies in the processing systems involved in the condition, but importantly, in other functional systems as well (Karmiloff-Smith, 2006). This research paradigm is in contrast to what has been called a modular approach, where only the superficial deficits associated with a condition are studied and selective developmental deficits (as selective to certain modules) are expected, even though it is only the mature brain and not the developing one which contains such functional specializations.

Research looking beyond the immediate phenotype (symptoms) of a condition has uncovered a number of unexpected and illuminating deficits in developmental disorders such as schizophrenia (Braff & Freedman, 2004), dyslexia (Ahissar, Lubin, Putter-Katz, & Banai, 2006), and autism (Sykes & Lamb, 2007). Each of which bring us closer to an understanding of how the genetic factors behind these conditions interact and produce the complex traits associated with these disorders. Furthermore, many of these discovered abnormalities can be used to aid genetic analysis as they are usually more simple (i.e. involve less genes) and robust than the standard phenotypic traits (symptoms) used to diagnose a condition. These discovered phenotypic traits are known as endophenotypes (Gottesman & Gould, 2003; Gottesman & Shields, 1972) and they are can be very informative, since they reside as intermediates between the neurodevelopmental condition and the genetic factors behind it. An endophenotype can be anything from a cognitive, neurophysiological or biochemical trait and to be the most useful it needs to be stable and inheritable. Recently, there has been a surge of interest in the concept of an endophenotype and its potential importance for understanding both the genetic and neurological underpinnings of neurodevelopmental disorders (see Cannon & Keller, 2006; Gottesman & Gould, 2003; Simon, 2007; Walters & Owen, 2007, for recent reviews)

Are there candidate endophenotypes associated with synaesthesia? Putative endophenotypes already exist in the synaesthesia literature. For example, synaesthetes have been linked to a greater incidence of a condition known as Mitempfindung (referral of a tactile sensation to a different location of the body than the point of tactile stimulation) (Burrack, Knoch, & Brugger, 2005), a greater memory for colours (Yaro & Ward, 2007), enhanced visual imagery (Barnett & Newell, 2007), creativity (Mulvenna, 2007) and artistic inclination (Ward, Thompson-Lake, Ely & Kaminski, 2007). Also recent neurophysiological data measuring evoked related potentials in linguistic-colour synaesthetes using stimuli that *do not* evoke synaesthesia, such as checker boards and Gabor stimuli, showed early differences in the visual evoked potential (in the C1 and P1 components) between synaesthetes and non-synaesthetes (Barnett *et al*., in preparation), indicating that synaesthetes have early visual sensory differences, possibly reflecting abnormal connectivity in these areas which may serve as an endophenotypic marker for synaesthesia.

Given that in synaesthesia two or more cortical or functional modules exhibit increased cross-talk, as a candidate endophenotype, the following set of experiments focused on cross-talk which is normally present between sensory systems i.e. the cross-talk involved in multisensory integration. Multisensory integration is where two or more perceptual modalities interact or combine to produce a distinct perceptual output that is distinct from either sense alone (Ernst & Bülthoff, 2004). An example of multisensory integration is the well known ventriloquism effect, where the speech coming from a ventriloquist is perceived as coming from the puppet's mouth rather than the ventriloquist's. This illusion (as with other multisensory effects) has been rigorously studied and verified in experimental settings with both complex and simple stimuli (Bertelson & Radeau, 1981; Howard & Templeton, 1966; Jack & Thurlow, 1973; Welch & Warren, 1980).

Multisensory integration abnormalities are thought to a play a role in other neurodevelopmental disorders. In autism, the sensory integration theory (Ayer, 1979) argues that autism is caused by a failure to develop normal multisensory connections. However, neurophysiological evidence for this theory is lacking (Molholm & Foxe, 2005) and behavioural measures have had contradictory results (Williams, Massaro, Peel, Bosseler & Suddendorf, 2004; Smith & Bennetto, 2007). The integration of visual information and phonetic information has been suggested to be a causal factor in dyslexia (Pammer & Vidyasagar, 2005), where recent evidence shows that the temporal window of integration in dyslexics for simple audiovisual integration is extended (Hairston, Burdette, Flowers, Frank, & Wallace, 2005; Foucher *et al.*, in press). Also, in schizophrenia patients, abnormal multisensory intergration of speech stimuli have been found (de Gelder, Vroomen, Annen, Masthof & Hodiamont, 2002).

It is surprising that in a neurodevelopment condition such as synaesthesia which involves increased cross-talk between cortical areas, that multisensory integration has not been examined. Also, investigating multisensory integration in synaesthetes may shed light on the hypothesis that synaesthesia is an enhanced form of multisensory integration, or a case of hyperbinding (Esterman *et al.*, 2006; Hubbard, 2007; Robertson, 2003). Evidence for this theory comes from the fact that non-synaesthetes show similar cross-modal associations as synaesthetes and recent TMS studies have shown that if parietal regions are deactivated with TMS pulses this abolishes the synaesthetic congruency effect (i.e. priming by synaesthetic colour) (Esterman *et al.*, 2006; Muggleton *et al.*, 2007). Also neuroimaging using fMRI

(Weiss *et al.*, 2005) and DTI (Rouw & Scholte, 2007) have both provided evidence that parietal regions may play a role in synaesthesia.

The study reported in this chapter examined a number of aspects of multisensory integration, with three different experiments. The first experiment used a simple audiovisual illusion known as the 'Shams' illusion or 'sound induced flash illusion' (Shams, Kamitani & Shimojo, 2000), to examine the degree of integration of simple visual and auditory stimuli, such as beeps and flashes. Experiment 2 examined the window of integration (how far apart in time the multisensory signals have to be in order for integration to occur) again using the 'Shams' illusion (Shams, Kamitani & Shimojo, 2002). The third experiment used more complex stimuli and also stimuli that induce synaesthesia i.e. speech sounds. This was accomplished using a well known audiovisual speech illusion called the 'McGurk effect'. In this illusion incongruent visual information influences the perceived auditory information (spoken word), resulting in a change in the perceived sound from the real auditory input (McGurk & MacDonald, 1976). In each of these experiments the primary research question was whether there are any differences between synaesthetes and non-synaesthetes in how information is integrated across the senses. For example, if synaesthetes perceived more multisensory illusions than non-synaesthetes, this would suggest that the cross-talk differences observed in synaesthesia are not localised to their synaesthesia but instead are more diffuse possibly indicative of a broader endophenotype.

2.1 Experiment 1

In order to examine cross-talk that is independent of the synaesthetes 'synaesthesia' a basic audiovisual integration was carried out. The task was based on a multisensory phenomenon known as the 'Shams' illusion, where a visual illusion is induced by sound (Shams *et al.*, 2000): when two short auditory beeps are presented together with a brief visual flash, 2 flashes are perceived most of the time. This illusory flash can occur with many different numbers of beeps and flashes and is the result of audition 'capturing' vision in this task (Shams, Ma & Beierholm, 2005). The explanation behind this phenomenon is that audition is more accurate at temporal judgements than vision and due to this the brain weighs information in favour of audition (if the task was a spatial task then the dominant modality would be vision). Explanations along this line have been worked out in detail with probabilistic models where the reliability of a signal in a multisensory task is directly related to its weight in a multisensory task; the more reliable a signal the more likely that it will capture or influence another competing sensory signal, and thus dictate the multisensory percept (Ernst, 2005; Ernst & Banks, 2002;). However, even though audition is more reliable at this task, the probabilistic models (e.g. maximum likelihood model of Ernst and Banks) also predict that in a smaller proportion of trials (the proportion is dependent on the differences in the reliability between the two signals) that the reverse illusion will also occur, that is, vision would capture audition thus inducing auditory illusions.

Also the type of illusion experienced can differ in that the participant can experience what are called fission illusions and fusion illusions (Andersen, Tiippana & Sams, 2004; Shams *et al.*, 2002). A fission illusion is when the pervceived number of events is greater in number than the actual number of stimuli. For example, if two flashes are reported upon presentation of one flash and two beeps,then this is called a flash fission response as the one flash 'breaks' into two. A fusion response is the opposite, in that the number of percieved events are fewer in number than the actual number of events in one modality. For example, using the same example as above, a beep fusion response would be when one beep is reported instead of two (due to one visual flash).

This means that by using the Shams illusion the degree of both auditory and visual capture experienced between synaesthetes and non-synaesthetes can be examined. To ensure that any differences observed in integration are not due to differences in the processing of the unimodal stimuli, information regarding the unimodal stimuli was required. For example, synaesthetes could show more visual capture than non-synaesthetes, but this may not be due to a difference in the integration process between synaesthetes and non-synaesthetes, instead it may simply reflect reliability differences in the processing of the unimodal signals. For example, synaesthetes maybe more reliable in dealing with visual information than non-synaesthetes. Also, individual differences could occur in the perception of the number of unisensory stimuli: a participant could experience two flashes presented in isolation as three flashes 50 percent of the time, so to get an accurate measure of how many illusions a participant experienced in an illusory condition (i.e. when there was a different numbers of flashes and beeps) this unimodal response would have to be compared with the illusory response.

A number of illusion conditions were carried out to eliminate any reporting bias that could be due to the repeated presentation of one illusory condition. It also allowed examination of the amount of integration across the groups when there was a small discrepancy between the sensory signals (one flash and two beeps) to when there was a larger discrepancy (one flash and three beeps). Different amounts of integration may occur depending on this discrepancy (e.g. the smaller the discrepancy the more integration) and it is possibile that the groups may show differential affects in integration across these different temporal discrepancies.

2.1.1 Methods

2.1.1.1 Participants. 12 synaesthetes (two male and 10 female) and 12 controls (sex and age matched) took part in the experiment and were compensated for their participation. The synaesthetes ranged in age from 18 to 61 years, with a mean age of 34.33 (S.D = 14.65), the controls ranged in age from 20 to 61 years, with a mean age of 35.5 (S.D = 15.25). All participants reported no hearing abnormalities and normal to corrected-to-normal vision. All synaesthetes reported experiencing grapheme-colour synaesthesia, and have been previously tested for consistency in their grapheme to colour associations (Barnett *et al.*, 2007). The study was approved by the School of Psychology Ethics Committee, Trinity College, Dublin. Informed, written consent was obtained from all participants prior to the experiment.

2.1.1.2 Apparatus and Stimuli. The visual (flash) stimulus consisted of a uniform white disc subtending 1.5° degrees of visual angle at an eccentricity of 10° below a fixation point, presented for a duration of 10 ms against a black background between one to three times. The auditory stimulus (beep) was 10 ms long with a frequency of 3.5 kHz and had a 1ms long ramp at either end of the sound wave envelope. The beep was presented either one to three times through speakers positioned at either side of the computer monitor and at the same height as the flashes (Shams *et al.*, 2005). Multiple flashes and beeps were presented with stimulus onset asynchronies (SOA's) of 70ms and 58ms respectively, where the relative timing between flashes and beeps was set such that the centre of the flash and beep sequences was synchronous, thus maximising the time overlap between the two stimuli (Shams *et al.*, 2005). Figure 2 illustrates the experimental setup and the temporal characteristics of the stimuli.

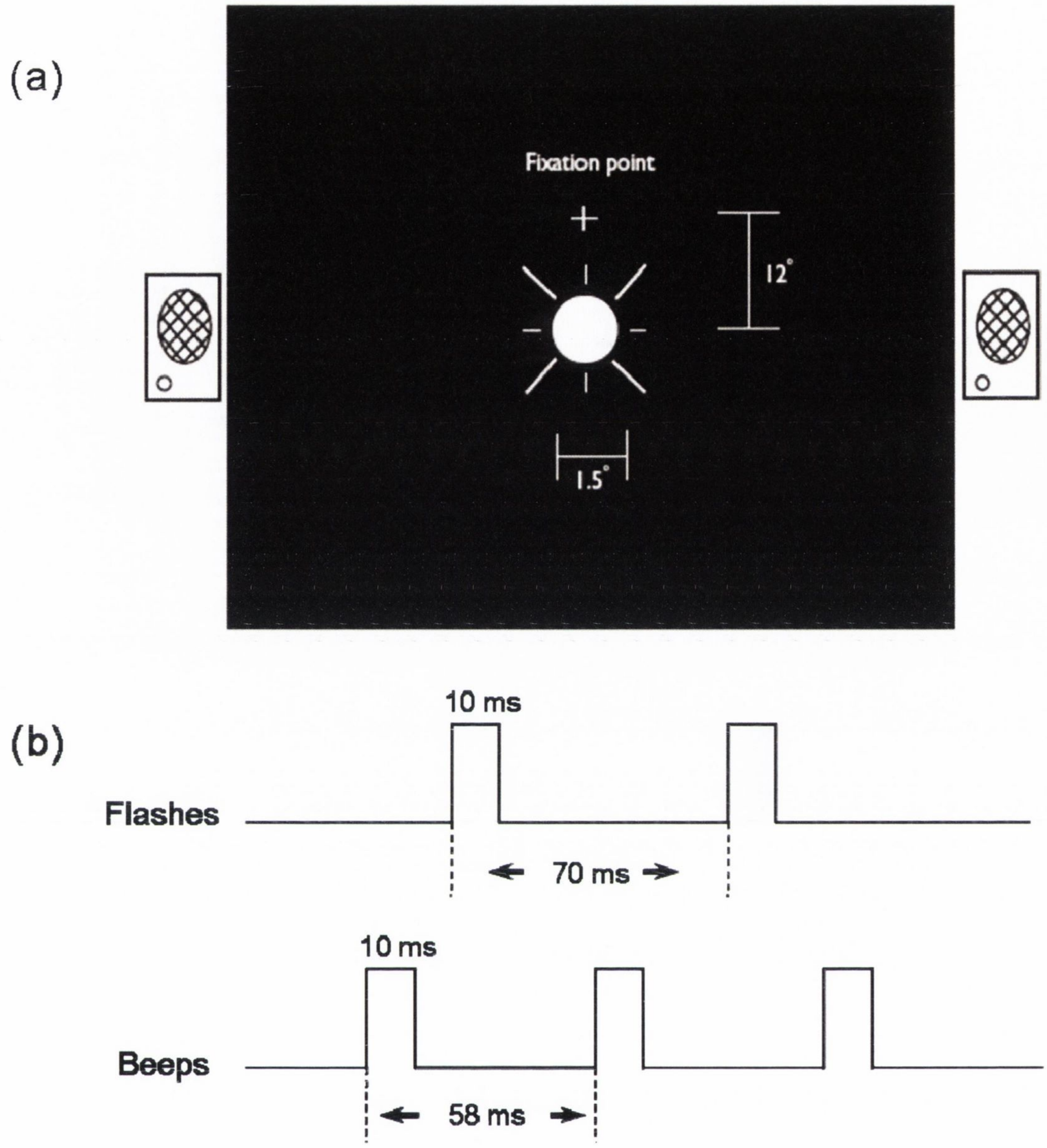

Figure 2: The apparatus and stimuli set-up used in Experiment 1.
(a) The speakers were positioned horizontally to the flash, and the flash was presented 10 degrees of visual angle peripheral to the fixation cross. (b) Multiple beeps and flashes were presented with temporal disparities of 70ms and 58ms for the flashes and beeps respectively, with each temporal sequence centred in relation to each other (as illustrated).

2.1.1.3 Design. The experiment was based on a mixed design with one between-subject factor of Group (synaesthetes versus non-synaesthetes) and one within-subjects factor with 15 conditions. The 15 different conditions comprised; three audio only conditions (one, two and three beeps), three visual only condition (one, two and three flashes), giving a total of six unimodal conditions. There were nine bimodal conditions containing different combinations of both beeps and flashes (one beep and one flash, one beep and two flashes, one beep and three flashes, two beeps and one flash, two beeps and two flashes, two beeps and three flashes, three beeps and one flash, three beeps and two flashes and three beeps and three flashes). This gave six conditions in which the visual or auditory sensory signals were not the same and in these conditions the discrepancy between the sensory signals resulted in either (in proportion of the trials) visual or auditory illusions. Each condition contained 20 trials, giving a total of 300 trials which were presented randomly, with a brief break every five trials. A practice block of 30 trials (two trials from each condition) was carried out before the experiment and data from these trials were excluded from the final analysis.

2.1.1.4 Procedure. Participants were tested individually in a windowless room with fluorescent lighting. They were seated in front of a computer monitor with their head resting on a chin-rest 57cm away from the monitor. Their task was to maintain fixation on a fixation point in the centre of screen (which remained on the screen throughout the whole experiment) and to indicate how many beeps or flashes occurred by pressing the corresponding number key on a keyboard. Each participant entered the numbers of beeps or flashes in the same order throughout the experiment; this order was counterbalanced among synaesthetes.

To measure the amount of illusions that occurred in each of the 6 illusion conditions (one flash and two beeps; one flash and three beeps; two flashes and three beeps; two flashes and one beep; three flashes and two beeps; three flashes and one beep), the responses (both flash and beep responses) in each of these conditions were compared with the unimodal condition for each stimulus. For example, if in the 'one flash' unimodal condition a participant reported two flashes 20 percent of the time, and then in the 'one flash and two beeps' illusory condition reported two flashes 50 percent of the time, the actual number of illusions was calculated as 30 percent of illusions (50 percent minus 20 percent). This procedure was carried out across all participants. Also, both the visual and auditory illusions were separated into the fission and fusion illusions for analysis, as fission and fusion illusions are thought to result from different underlying processes (Shams *et al.*, 2002).

2.1.2 Results and Discussion

One of the synaesthetes could not discriminate one from either two or three beeps (in any trial) and thus their data were removed from all analysis. The percentage of illusions across each of the six illusory conditions was calculated for each synaesthete taking into the account the unimodal responses. To assess if synaesthetes differed from non-synaesthetes in the amount of auditory capture (i.e visual illusions) a 2 by 6 mixed ANOVA was carried across the six different illusory conditions, incorporating both flash fission and flash fusion illusions. There was no significant main effects of condition, $F(5, 100) = 1.118$, $p = 0.35$, or of group, $F(1, 20) = < 1$ found. Critically there was no interaction between these factors, $F(5, 100) = 0.767$, $p = 0.523$. The same analysis was carried out on the beep illusions (visual capture) across the six conditions and a mixed ANOVA revealed no significant effect of condition, $F(5, 100) < 1$, group, $F(1, 20) < 1$, or interaction, $F(5, 100) < 1$.

A further analysis of the amount of fission illusions (when the participant reports *more* beeps or flashes that are present) and fusion illusion (when the participant reports *less* beeps or flashes that are present) responses for both flash and beep illusions was carried out across conditions (see Figure 3). A 2 (group) by 4 (flash fission, flash fusion, beep fission, beep fusion) mixed ANOVA was carried out and revealed a significant effect of condition, $F(3, 60) = 18.017$, $p < .001$. There was no effect of group, $F(3, 60) < 1$ and no interaction, $F(1, 20) < 1$. The significant effect of condition was not unexpected since audition is more reliable than vision in this task (Shams *et al.*, 2005) and thus more flash illusions should be experienced. This is confirmed by examining the unimodal beep and flash variances depicted in Figure 4. A 2 by 2 mixed ANOVA confirmed that the variances were different between each modality, $F(1, 20) = 30.175$, $p < .001$, and there was no significant effect of group, $F(1, 20) < 1$, and also no interaction, $F(1, 20) < 1$. Given that the variance of the unimodal stimuli was the same across groups, this rules out the possibility that there were undetected differences in the integration process between the synaesthetes and non-synaesthetes. Synaesthetes may have had differences in unimodal sensory processing (i.e. synaesthetes may exhibit more reliable visual response) resulting in differential reliabilities between the two signals than controls, and as they showed the same amount of visual and auditory capture as controls, this would have of indicated that the two groups integrate information differently.

In order to confirm that there were no differences between groups in the proportion of illusions that were either visual or auditory a further analysis was carried out. The dependent variable was the proportion of visual illusions in each of the six illusory conditions. Using a 2 (group) by 6 (conditions) mixed ANOVA, no significant effect was found of condition, $F(5, 100) = 1.937$, $p = 0.439$, or group, $F(1, 20) = 0.862$, $p = 0.364$, also no interaction was observed, $F(5, 100) = 0.335$, $p = 0.818$.

These analyses indicate that synaesthetes and non-synaesthetes did not differ in the amount of visual or auditory capture induced by simple stimuli. It also revealed that there were no differences between each group in the processing of the unimodal stimuli used. This would suggest that the factors (genetic and otherwise) that contribute to synaesthesia are specific and that, at the very least, these factors do not contribute to cross-talk between simple audiovisual stimuli.

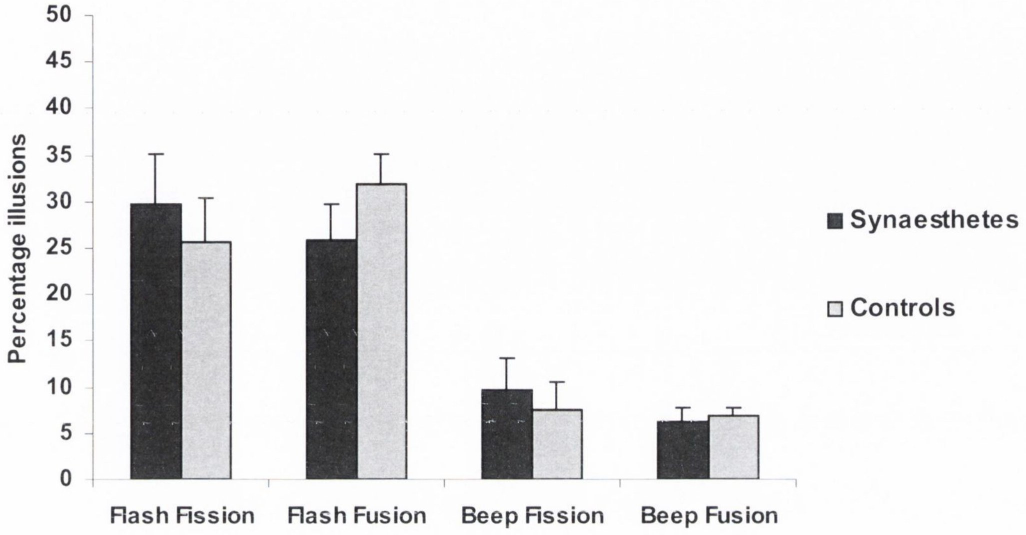

Figure 3. Plot illustrating percentage of illusions for both flash and beep illusions (fission and fusion illusions) across synaesthetes and controls (Experiment 1).

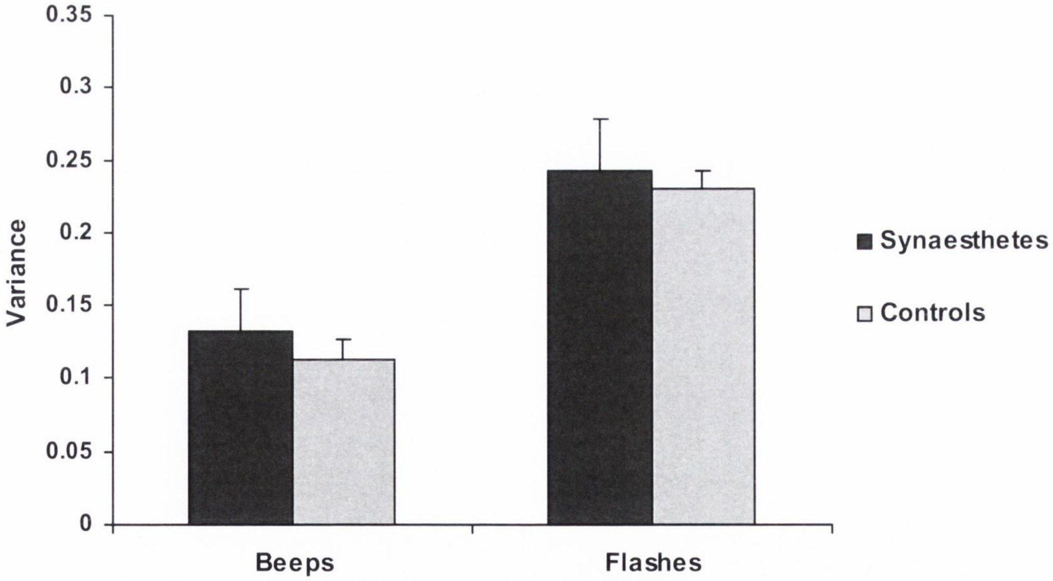

Figure 4: Plot illustrating unimodal variances (sd^2) for the all the beep and flash responses across synaesthetes and controls (Experiment 1).

2.2 Experiment 2

The degree of influence one sensory signal has on another or has on the integration process is not simply dependent on the relative reliability of the sensory signals, other factors such as spatial or temporal proximity have also an important influence. Indeed, two of the three principles of integration at the neuronal level, as outlined by Stein and Meredith (1993), are the temporal and spatial proximity of the stimuli. The multisensory neural response is greater when signals are close together in space and in time (Stein & Meredith, 1993). This is also been reflected in behavioural performance (Holmes & Spence, 2005) in that integration occurs to a greater degree when the stimuli are closer together in time and space and can disappear altogether when the spatial and temporal discrepancies are large (Jack & Thurlow, 1973; Jones & Jarick, 2006).

The 'Shams' illusion is highly susceptible to temporal discrepancies and the proportion of flash illusions drops off rapidly as the beeps become further apart in time i.e. 100-150 ms (Shams *et al.*, 2002). This indicates that there is a certain temporal window during which integration of simple stimuli occurs and that it is approximately 200-300ms in size. This temporal window of integration has been examined for a number of different tasks using different types of multisensory phenomenon and with both simple and complex stimuli (Horváth, Czigler, Winkler & Teder-Sälejärvi, 2006: Jones & Jarick, 2006; Wang, Datta, & Sussman, 2005). It has been shown that, in concordance with the temporal window of integration in the Sham's illusion, introducing temporal discrepancies of greater than 100ms between audio and visual stimuli results in a significant decrease in integration in other multisensory phenomenon, such as the ventriloquist effect (Lewald, Ehrenstein & Guski, 2001; Lewald & Guski, 2003; Radeau & Bertelson, 1987; Slutsky &

Recanzone, 2001; Thomas, 1941). When using more complex meaningful stimuli the temporal window of integration is typically extended (Jack & Thurlow, 1973).

This principle of multisensory integration has been examined in other neurodevelopmental conditions such as dyslexia (Hairston, Burdette, Flowers, Wood & Wallace, 2005; Foucher *et al.*, in press). Hairston *et al.* (2006) showed that dyslexics integrate simple audiovisual stimuli over a larger temporal window than non-dyslexics in a task involving a visual temporal order judgement with concurrent task-irrelevant auditory information of different stimulus onset asynchronies (to the second visual flash). Likewise, it is also possible in synaesthetes that the same amount of audiovisual integration is found as controls (as Experiment 1 indicates), but that they may integrate over a wider temporal window than non-synaesthetes. Such a result would suggest that the processes that mediate integration are functionally different in synaesthetes.

2.2.1 Methods

2.2.1.1 Participants. Eleven synaesthetes (all female) and 11 controls (sex and age matched) took part in this experiment and were compensated for their participation. The same synaesthetes and controls which performed the first experiment also participated in this experiment with the exception of one synaesthete (and this synaesthetes matched control) who did not partake in this second experiment. Experiment 1 and 2 were carried out in the same session, with a ten minutes break in between both experiments, each took approximately 40 mins. The synaesthetes ranged in age from 18-61 years, with a mean age of 34.55 (S.D = 15.36), the controls ranged in age from 20-61 years, with a mean age of 36 (S.D = 15.88). All participants reported no hearing abnormalities and normal or corrected-to-normal vision. All synaesthetes reported grapheme-colour synaesthesia and have been previously tested for consistency in their grapheme to colour associations (Barnett *et al.*, 2007). The study was approved by the School of Psychology Ethics Committee, Trinity College, Dublin. Informed, written consent was obtained from all participants prior to the experiment.

2.2.1.2 Stimuli. The stimuli configuration was identical to the previous experiment, differing only in how many beeps and flashes occured. The visual stimulus was always one flash (except for one condition of catch trials where two flashes were presented), whereas the auditory stimulus was always two beeps, one beep of which was presented at the same time as the flash whereas the other beep was presented at one of the following stimulus onset asynchronies (SOA); 50, 70, 90, 110, 130, 150, 170, 190, 210, 230 and 250 msec after the flash and SOA's before the flash; -50, -70, -90, -110, -130, -150, -170, -190, -210, -230 and -250 msec. In the catch trials two flashes and two beeps were presented, the flashes were separated by 70 ms and the beeps by 58 milliseconds. This catch trial condition was included to prevent a reporting bias.

2.2.1.3 Design. The experiment was based on a mixed design with a between-subject factor of group (synaesthetes versus non-synaesthetes) and a within-subjects factor of 23 conditions. The 23 conditions contained 11 positive SOAs, 11 negative SOAs and one condition containing 2 flashes and 2 beeps (i.e. catch trial condition). Each condition contained 15 trials, giving a total of 345 trials which were presented randomly, with a brief, self-timed break every few trials. For all participants this experiment was carried out immediately (following a short break) after the previous experiment (Experiment 1).

2.2.1.4 Procedure. The procedure was identical to the previous experiment, except that the participants only had to report (by pressing the corresponding number key) how many flashes occurred (and not how many beeps were heard).

2.2.2 Results and Discussion

Synaesthetes who did not get more than 50 % of the catch trials correct were excluded from the data analysis; this resulted in excluding the data from three synaesthetes. Accordingly, data from their matched controls were also excluded. The proportion of visual illusions at each of the temporal of the SOA's for both groups is presented in Figure 5.

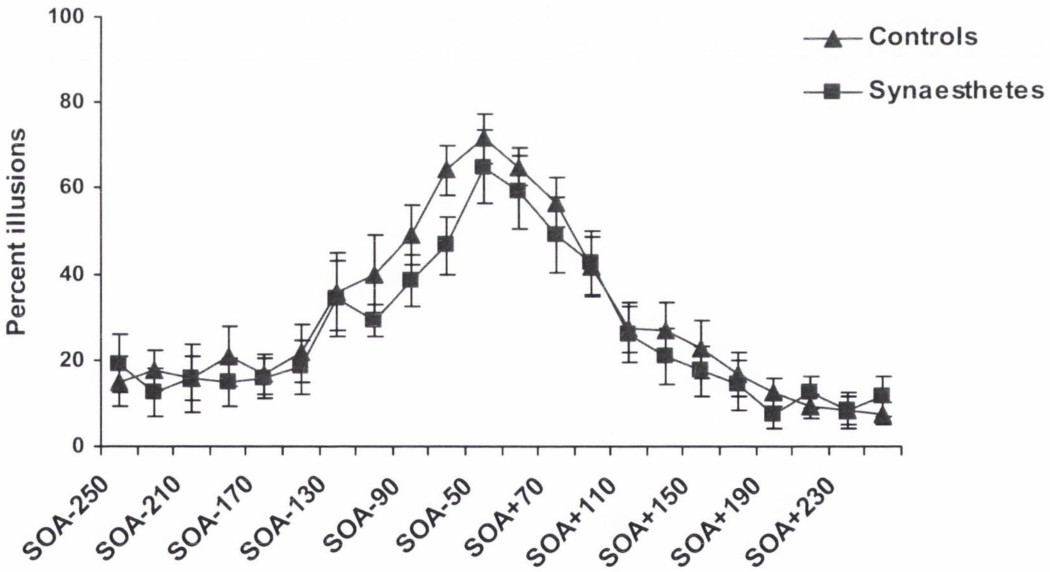

Figure 5: Plot illustrating percentage of illusory responses reported across different stimulus onset asynchronies between two beeps, for both synaesthetes and controls (Experiment 2). One beep is concurrent with a flash and the other was presented either before or after this beep. Data are shown separately across each group.

.

The temporal profile across the SOAs was similar across groups the peak percentage of illusions occurred in the same SOA, this indicates that synaesthetes and controls did not have shifted temporal profiles (in relation to each other) i.e.

synaesthetes and controls both integrated maximally when the second beep was close to simultaneous with the flash. Given this similarity in the temporal profiles across groups, the next step was to determine whether different proportion of illusions occurred overall across the entire temporal profile. To measure this the area under the curve was chosen, this calculation enables an overall temporal measure of the proportion of illusions, e.g. synaesthetes or controls may not show significant differences at any one SOA but over all the SOA's, the area under the curve might show a different overall pattern in the temporal profile between the groups. The area under the curve was calculated for both groups using the trapezoid rule. No significant difference was found between the mean area under the curve across the groups, $t(14) = 0.475$, $p = 0.642$, with mean values of 100.33 for synaesthetes and 116.58 for controls. Also, independent comparisons between the groups at each SOA were carried out using a modified Bonferroni (Bonferroni-Sidak adjustment (Sidak, 1967), see Keppel and Wickens (2004) for a discussion of this correction) and did not reveal any significant differences. This indicates that synaesthetes and non-synaesthetes integrate information similarly across different temporal discrepancies for simple stimuli such as beeps and flashes, however at the –100 SOA there is approximately a 20% difference in the proportion of illusions but this did not prove significant. A larger sample of synaesthetes (given the high exclusion rate) may have been useful here. Also, as in Experiment 1, there was no evidence that would indicate that synaesthetes experience more visual illusions in general than non-synaesthetes. The performance of the synaesthetes appeared to be identical to that of non-synaesthetes not only in the amount of integration between the senses but also in how this integration is effected by other factors such as temporal proximity.

2.3 Experiment 3

Experiment 1 and 2 suggest that synaesthetes integrate simple audiovisual stimuli in a similar fashion as non-synaesthetes. However this may not hold up for more complex audiovisual stimuli which would involve different cortical areas or for stimuli that induce synaesthesia. For example, in an investigation of multisensory processing in a schizophrenic population (de Gelder, Vroomen, Annen, Masthof & Hodiamont, 2002), it was found that for simple stimuli such as beeps and flashes normal integration occurred but for more complex stimuli (i.e. speech sounds, using the McGurk effect (see below)) diminished multisensory integration was observed in schizophrenics compared to controls. Also, Molholm & Foxe, (2005) reported that schizophrenics benefited less from visual articulations of speech in noisy acoustic conditions than controls. Given that synaesthesia is so commonly associated with linguistic factors (Simner, 2007) any differences in cross-talk in synaesthetes versus non-synaesthetes may be confined to stimuli which are related to their synaesthesia, i.e. spoken words or graphemes. The closer the cortical areas are to the locus of the synaesthetic abnormal connectivity (structural or functional) the more likely it will be that any other related effects may be observed. We considered examining multisensory integration for graphemes and phonemes but choose spoken words instead, given that a well studied illusion (known as the McGurk effect (see below)) exists for spoken words. Also the perceptual status of mulitsensory integration for phonemes and graphemes is controversial (in that it can be attributed to mere memory association).

The McGurk effect (McGurk & MacDonald, 1976)is an audiovisual speech illusion where incongruent visual information influences the perceived auditory information, resulting in a change in the perceived sound from the real auditory input.

In their original paper, McGurk and MacDonald used the following combination of stimuli to induce the illusion: they presented an audio of an actor saying /ba/ and dubbed this onto a viseme of an actor lip-saying [ga]. Participants are then asked to report what was heard and, in the majority of cases, an observer perceives an illusory syllable "da", that is, a fusion of the two inputs. In the following study, instead of using syllables words were used, e.g. /bait/ + [gate] = "date" (Alsius, Navarra, Campbell & Soto-Faraco, 2005), which enabled the presentation of many different illusory trials.

2.3.1 Methods

2.3.1.1 Participants. Twelve synaesthetes (one male and 11 female) and 12 controls (sex and age matched) took part in the experiment and were compensated for their participation. Experiment 3 was carried out on a different set of synaesthetes than in the synaesthetes in Experiment 1 and 2. The synaesthetes ranged in age from 24 to 63 years, with a mean age of 43.9 years (S.D = 13.43). The controls ranged in age from 20-61 years, with a mean age of 41.7 years (S.D = 14.02). All participants reported no hearing abnormalities and normal or corrected-to-normal vision. All synaesthetes reported grapheme-colour synaesthesia, and were previously tested for consistency in their grapheme to colour associations (Barnett *et al*., 2007). The study was approved by the School of Psychology Ethics Committee, Trinity College, Dublin. Informed, written consent was obtained from all participants prior to the experiment.

2.3.1.2 Stimuli and apparatus. Visual and audio recordings from a male speaker (clean shaven, aged 25, full face view) were used to create the stimuli. A camcorder (JVC digital video camera, model: GR-DVL167) recorded both video and audio samples. The actor was instructed to speak certain words at a normal pace and each video clip was two seconds long. These words were selected based on phonetic properties known to give rise to McGurk illusions when artificially dubbed (Alsius *et al.*, 2005; McGurk & MacDonald, 1976).

Each visual-speech video clip was edited using professional video editing software (Adobe Premiere 6.0) for the PC. From each of the audio and visual stimuli, a basic set of audiovisual stimuli (AV) were first created by pairing these stimuli to create with incongruent AV pairs (i.e. McGurk stimuli) with 32 stimuli in this set. To create the incongruent AV (or McGurk) stimuli, the audio version of one word with the viseme of another word were spliced together. Previously known properties of phonemes and their articulations were used to generate McGurk illusions (Alsius *et al.*, 2005; McGurk & MacDonald, 1976). More specifically,audioand visual pairings were conducted in a non-random manner and were constrained by the place of articulation of a critical phoneme: this critical phoneme had a more posterior place of articulation for the viseme (e.g. /g/, /k/, /n/) than the audio word (e.g. /b/, /p/, /m/, respectively) (Alsius *et al.*, 2005; McGurk & MacDonald, 1976). Some combinations of words produced an entirely new word as in /bait/ + [gate] = 'date', while in other combinations the McGurk fusion matched with the visual word e.g. /bent/ + [dent] = 'dent' (see Table 1 for a list of these AV pairings and the expected illusory response).

A set of Visual-only (V) and Auditory-only (A) stimuli were also created and matched each of the A and V components of the incongruent AV set of stimuli. Visual-only stimuli were created by masking the incongruent audio component (i.e. the actor's spoken word) with white noise. To create the Audio-only stimuli a previously reported technique was adopted (Alsius *et al.*, 2005; Campbell & Massaro, 1997; MacDonald, Andersen & Bachmann, 2000) in that the incongruent viseme was masked using spatial quantization with the degree of pixelation measuring 10 pixels horizontally across and 15 pixels vertically along the face image (see Figure 6). Each masking procedure was employed to disrupt recognition of the masked unimodal stimulus with a minimum disruption to the amount of information in each condition (Alsius *et al.*, 2005).

Table 1: List of experimental word stimuli and corresponding expected percepts used in Experiment 3

Audio Word	Visual Word	Expected McGurk Word	Audio Word	Visual Word	Expected McGurk Word
BEEN	BEEP	BEAM	LISP	LIST	LIST
CAP	CAN	CAT	NAIL	PALE	MAIL
COP	CON	COT	MAP	MAT	MAT
PRAM	CRAM	CRAM	NAY	PAY	MAY
BAIT	GATE	DATE	NEET	PEAT	MEET
BAY	GAY	DAY	MAIL	NAIL	NAIL
VET	DEBT	DEBT	MAP	NAP	NAP
VET	GET	DEBT	MAP	NAP	NAP
VEER	DEER	DEER	MAY	NAY	NAY
VEER	GEAR	DEER	MICK	NICK	NICK
BENT	DENT	DENT	RAN	RAP	RAM
BOG	DOG	DOG	RIB	RIG	RID
GRIN	GRIP	GRIM	SHOP	SHOCK	SHOCK
PALE	NAIL	TAIL	PALE	TAIL	TAIL
PIP	NIP	TIP	PIN	TIN	TIN
HIP	HIT	HIT	WARN	WARP	WARM

The experiment was programmed and all stimuli were displayed using Presentation® software on a PC. Participants were tested individually in a windowless room with fluorescent lighting. They were seated at a distance of 57cm from the monitor with the result that the visual image of the AV display subtended a visual angle of 18 degrees.

2.3.1.3 Design. The experiment was based on a two-way, mixed design with three different levels: A, V and AV (incongruent). Trials in each condition were presented in separate blocks and the order of the blocks was counterbalanced across participants according to a latin square design. Each block contained 32 trials, the first two of which were practice trials (and the responses were excluded from the data analysis) and the remaining 30 were presented in random order across participants.

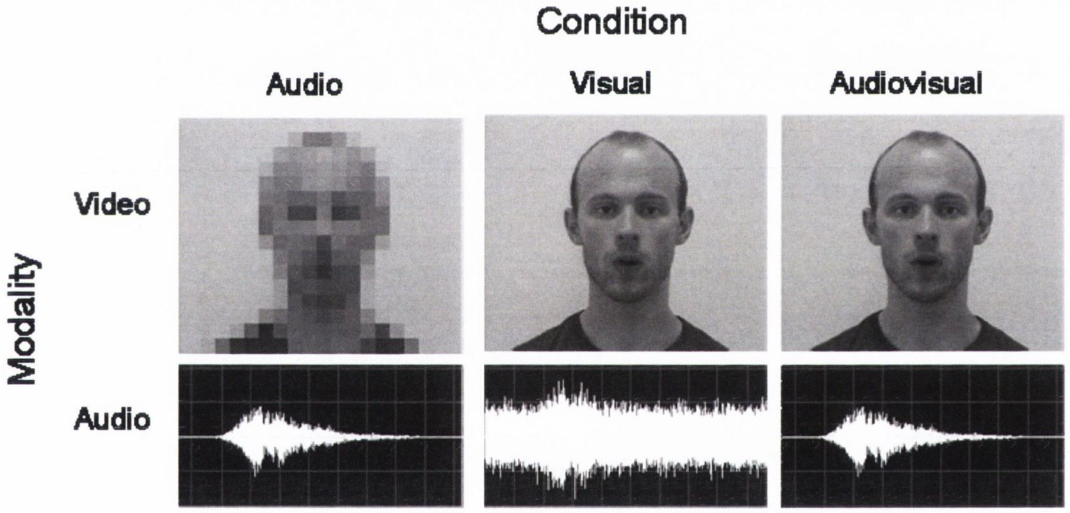

Figure 6: The visual and audio component of each audiovisual speech stimulus for each condition (Experiment 3 & 4).
For the A-only condition, a spatial quantization technique was used to mask the visual information; likewise in the V condition, white noise was used to mask the auditory signal. Each masking procedure was employed to disrupt recognition of the masked unimodal stimulus with a minimum disruption to the amount of information in each condition.

2.3.1.4 Procedure. A trial consisted of a fixation cross which was presented in centre of the screen for 1000ms. Following fixation, either an A, V or AV (incongruent) stimulus appeared for 2000ms. Participants were instructed to determine what was being spoken by the actor in each trial as accurately as possible (and if in doubt, to give their best guess). However, in order to investigate a different research question (see Chapter 3), prior to reporting what they heard, the participant also picked a colour from a colour picker. Synaesthetes were requested to choose the colour that best matched their synaesthetic colour experienced was chosen whereas controls were requested to choose any colour that came to mind. The colour picker was presented on a different monitor which was positioned to the right of the participant (see Chapter 3).

The verbal responses of the synaesthetes were recorded by the experimenter (the experimenter was aware of the group and condition), and were later classified as either illusory or not following previously reported criteria (Alsius *et al.*, 2005; Windman, 2004) Specifically, the illusion had occurred when the result was a novel fused percept (not the same as either of the inputs) or when the auditory input had been changed to match the visual component (viseme). Following Alsius *et al.*, (2005) each response was then classified into either a visual/fusion response (i.e. a McGurk illusion), an audio response (when the response corresponded to the audio component) or some other response unrelated to any component, which was labelled as ‘other’. Table 1 was used to determine which category each response fell into (see 2.3.1.2).

2.3.2 Results and Discussion

All of the synaesthetes and non-synaesthetic controls were susceptible to the McGurk illusion. To examine whether synaesthetes and non-synaesthetes differed in their susceptibility to the McGurk illusion, the proportion of visual or fusion responses (classified as a visual/fusion response, i.e. McGurk illusions) that occurred were determined for each of the three stimulus conditions; Audio alone, Visual alone and audiovisual. The proportion of McGurk illusions reported in each condition was the dependent variable of interest and was subjected to a 2 by 3 mixed ANOVA, with one between subject factor (i.e. group) and three within-subjects factors (A, V and AV). See Figure 7 for the mean number of illusions reported in each condition.

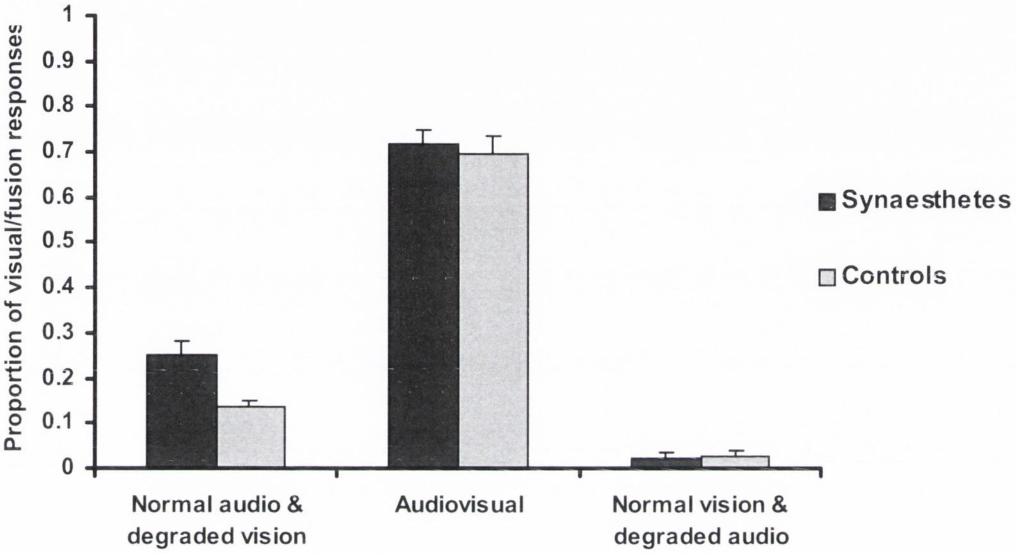

Figure 7: Plot illustrating proportion of visual or fusion response across each condition (Experiment 3).

A main effect of condition was found, *F*(2, 44) = 415.702, $p < 0.001$, which was due to the larger number of illusions experienced in the AV condition than either the A or V conditions. This was to be expected, given the effectiveness of the masking procedures in the unimodal conditions. No effect of group was found, *F*(1, 22) = 3.695, $p = 0.068$, although to the effect did approach significance. There was, however, a significant interaction between the factors, *F*(2, 44) = 3.393, $p < 0.05$. Post-hoc analysis, with Sidak adjustment for multiple comparisons, revealed the source of the interaction to be a significant difference between the groups in the proportion of illusions reported in the audio alone condition, *t*(22) = 3.736, $p < 0.01$, with mean values of 0.25 for synaesthetes and 0.14 for controls. There was no significant difference between groups in the other two conditions, $t(22) = 0.371$, $p = 0.714$, and, $t(22) = 0.264$, $p = 0.794$, for the audiovisual and visual alone respectively. This significant difference between groups in the audio alone condition was unexpected since the masking procedure was used to effectively reduce input from the visual signal. However, previous research has shown that the McGurk illusion can occur if the pixelation (caused by the spatial quantization technique) is as course as 11.2 pixels horizontally across the face (MacDonald *et al.*, 2000). The pixelation in this study was 10 pixels horizontally across the face and was identical to a previous study also using (Spanish) words (Alsius *et al.*, 2005) and which did report some integration in the audio alone condition but not as much as observed here (i.e. their proportion of illusions in the visual condition in their first experiment (there were two different tasks) were 0.03 and 0.05 and for their second experiment were 0.09 and 0.12.

The lack of any significant difference between groups in the audiovisual condition may be down to both groups performing optimal integration in this condition. In line with this idea are results from an earlier pilot study, which showed that when the audio is presented on its own (without masked visual information) the proportion of the audio stimuli correctly identified was 0.71, suggesting that at similar proportions for integrating information performance may have been close to optimal (here performance for the synaesthetes and controls was at 0.72 and 0.70, respectively). In the audio alone condition, on the other hand, the visual mask may have allowed a small amount of effective visual information through, possibly resulting in performance that was at the limits of audiovisual speech integration. It is possible that under such conditions any differences in integration would become more readily detectable. Thus, in noisy (visual) environments this result suggests that synaesthetes are likely to perceive speech more accurately than non-synaesthetes.

Also this result is not down to differences in the variability in the unimodal responses as the error (or variability), as measured by the proportion of responses that were neither the audio, illusory or visual component (classified as proportion of 'other' responses), was not different across the groups. The proportion of 'other' responses was subjected to a 2 by 3 mixed ANOVA, and no significant effect of group, $F(1, 22) = 0.485$, $p = 0.494$, or interaction was observed, $F(2, 44) = 0.452$, $p = 0.639$. There was a main effect of condition, $F(2, 44) = 1602.776$, $p < 0.001$, but this simply reflects the large of amount of 'other' responses in the visual alone condition, 0.96 for synaesthetes and 0.97 for controls. An analysis of the audio responses across the 3 conditions further confirmed that the integration differences observed reflected integration rather than unisensory differences: a 2 by 3 mixed ANOVA revealed a significant effect of group, $F(1, 22) = 7.747$, $p < 0.05$, and an interaction, $F(2, 44) = 5.628$, $p < 0.01$, also a significant effect of condition was observed, $F(2, 44) = 470.945$, $p < 0.001$, (as was expected). Post hoc analysis, with Sidak adjustment for multiple comparisons, revealed that the source of the interaction was again a difference between the groups in the audio alone condition $t(22) = 3.562$, $p < 0.01$, with mean values of 0.62 for synaesthetes and 0.77 for controls (see Figure 8). This result shows that an increase in integration in the audio alone condition for synaesthetes also resulted in a decrease in correct audio responses and vice versa for controls, rather than different amounts of 'other' responses occurring which would indicate different amounts of error or variability between the groups.

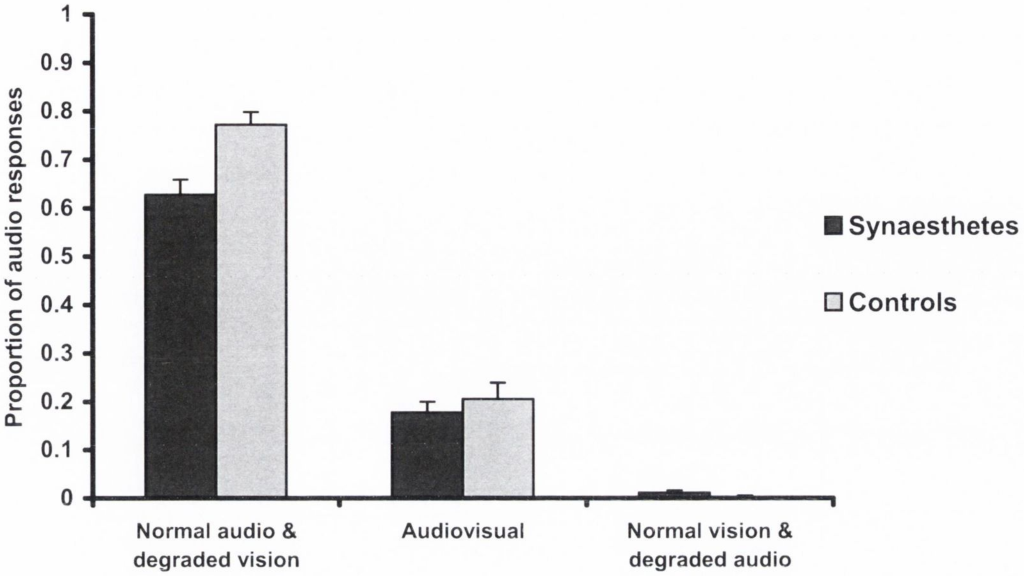

Figure 8: Plot illustrating proportion of correct audio response across each condition (Experiment 3).

2.4 General discussion

The three experiments reported here examined various aspects of audiovisual integration in synaesthetes and found that synaesthetes integrate information for simple stimuli (beeps and flashes) in a similar fashion to non-synaesthetes, in both the degree of 'capture' (visual or auditory capture) and the extent of the integration across different temporal disparities. However, for more complex stimuli, such as speech sounds, synaesthetes exhibited enhanced integration when substantial noise existed in one of the sensory signals (visual).

Experiment 1 was designed to investigate the quantity of auditory and visual capture using the so-called 'Shams' illusion. No difference was found between synaesthetes and non-synaesthetic controls in the amount of auditory and visual illusions or in the relative proportion of each. This result suggests that whatever factors contribute to synaesthesia they do not appear to affect cross-talk between visual and auditory areas in general or in respect to low level auditory or visual areas, as the Shams illusion is thought to involve processing between such areas (Bhattacharya, Shams & Shimojo, 2002; Mishra, Martinez, Sejnowski & Hillyard, 2007; Shams, Kamitani, Thompson & Shimojo, 2001; Watkins, Shams, Josephs & Rees, 2007; Watkins, Shams, Tanaka, Haynes & Rees, 2006). Experiment 2 was designed to investigate the amount of integration (auditory capture) that occurred over different acoustic temporal disparities. By presenting the auditory stimuli at different temporal disparities a profile of a temporal window of integration could be determined for these basic stimuli. The synaesthetes and non-synaesthetic controls had a similar temporal profiles, indicating that synaesthetes are constrained in the same way by temporal factors in audiovisual integration as non-synaesthetes, providing further evidence that the extra cross-talk observed in synaesthesia may be localised to areas or modules specifically involved in their synaesthesia.

The results of Experiment 3, however, revealed differences in synaesthetes when integrating speech sounds using the McGurk effect. These differences were only observed in the audio alone condition, where there was still a small amount of visual information available after the masking procedure to facilitate audiovisual integration. In such a scenario integration became more difficult and most likely amplified differences between synaesthetes and controls i.e. noise has less of an effect in hindering integration in synaesthetes than controls. This would suggest that in synaesthetes increased cross-talk exists between the cortical areas involved in the McGurk illusion which facilitate such enhancement of integration when the signal to noise ratio is decreased. Interestingly, the reverse effect (i.e. decreased integration) was found in patients with schizophrenia when the noise level was increased, the noise in this study was acoustic rather than visual (Molholm & Foxe, 2005) suggesting that, when the integration processes are pushed to their limit differences between groups are more readily visible.

Taken together these results provide evidence for a broader or possible endophenotype in synaesthesia. It is reasonable to assume that the genetic and developmental factors that contribute to the phenomenon of synaesthesia also contribute to this enhancement in audiovisual speech integration, either through affecting the structural or functional connectivity of the cortical areas involved. Indeed, a recent DTI study provided evidence of structural connectivity differences in the brain of synaesthetes in the right fusiform gyrus, left intraparietal sulcus and frontal areas (Rouw & Scholte, 2007). A possibility, not yet explored, is that a proportion of this extra connectivity may reflect broader connectivity differences unrelated to their synaesthesia (but not independent of the causal factors behind synaesthesia), and may even be involved in the increased audiovisual speech integration observed in this study. Interestingly, Rouw and Scholte (2007) found that the nature of the synaesthetic experience as measured by a questionnaire (whether the synaesthete experienced the synaesthesia in the minds eye – associator, or in the world - projector) only correlated with the increased connectivity in the right fusiform gyrus but not in the other cortical areas with increased connectivity. This does not rule out, of course, the possibility of the involvement of parietal and frontal regions in synaesthesia, especially given recent evidence from neuroimaging (Weiss *et al.*, 2005) and TMS studies (Esterman *et al.*, 2006; Muggleton *et al.*, 2007), for their involvement. Also the finding from this study suggests that heteromodal areas (Calvert, Campbell & Brammer, 2000; but c.f., Saint-Amour, De Sanctis, Molholm, Ritter & Foxe, 2006) involved in integrating audiovisual speech information are affected by the causal factors behind synaesthesia.

The exact role of these heteromodal areas in synaesthesia have been suggested to be in 'binding' features (Esterman *et al.*, 2006; Mulvenna & Walsh, 2006; Robertson, 2003; Weiss *et al.*, 2005). The leading proponent of the cross-activation model has recently argued that grapheme-colour synaesthesia may result from a two stage process; an initial early cross-activation in the fusiform gyrus followed by an upstream 'hyperbinding' mechanism in the IPS (Hubbard, 2007). The results presented here are not in agreement with an all-purpose 'binding' mechanism or a 'hyperbinding' mechanism in synaesthetes. This study demonstrates that synaesthetes 'bind' basic stimuli in a similar fashion to non-synaesthetes, but yet show an enhancement with audiovisual speech stimuli, suggesting that these processes are dependent on different neural systems. Furthermore, the idea of an upstream 'binding' module is also not in agreement with the recent multisensory literature which show that from the earliest stages of unisensory processing integration is occurring (see Schroeder & Foxe, 2005; Ghazanfar & Schroeder, 2006, for recent reviews), which can be facilitated by direct feedforward connections from other unisensory areas (Foxe & Schroeder, 2005). This highlights that interactions between the senses are distributed and involve many independent systems i.e. there is no requirement for a place in the brain 'where it all comes together' (Dennett, 1991).

The results of this study lead to the question of why audiovisual speech stimuli show enhanced integration whereas basic stimuli such as beeps and flashes do not. We might ask what is special about these stimuli? This may be related to the more fundamental question in synaesthesia research of why linguistic elements are so often the inducers in synaesthesia (Simner, 2007). Both these questions become more pertinent when the evidence for a single developmental mechanism uniting all forms of synaesthesia is considered, as it suggests that broad connectivity differences (or at the very least potential connectivity differences) exist in developing brain of synaesthetes and that some other factors dictate which type of synaesthesia develops. More often than not the type of synaesthesia that does develop involves language.

There are three types of explanations that may explain this trend; a genetic, developmental and a cognitive/informational. A simple genetic explanation would be that the genetic factors behind synaesthesia only have affects in areas of the cortex involved in language. A developmental explanation may involve a consideration of the developmental trajectory of the cortical areas (proto-cortical areas) involved in language and other sensory processing exhibit certain interactions which increase the probability that extra connections between these areas would be retained into adulthood. The final type of explanation possible, related to the latter idea, would be that informational or cognitive factors which are encountered when learning symbols may facilitate the consolidation of abnormal cross-modal interactions in the brains of developing synaesthetes. For example, learning to read and thus learning how to link graphemes to phonetic information involves a complex multisensory task (Pammer & Vidyasagar, 2005) linking perceptual features to language related phenomenon. If this process was overactive due to diffuse connectivity in the brains of synaesthetes and other perceptual basic categories (i.e. colour) were tagged on they might get consolidated, given the repetitive, explicit and conscious learning involved (unlike many other basic multisensory processes where this kind of explicit learning does not take place). Recently it has been suggested that synaesthesia may be an exaggeration on the processes that link perception and language (Simner & Ward, 2006), the results of this study would agree with such an interpretation. Clearly, any complete model of synaesthesia is going to have to explain why language plays such a prominent role.

In conclusion, this study found enhanced audiovisual speech integration in synaesthetes, whereas no such enhancement (or deterioration) was found with basic audiovisual stimuli. This suggests that broader cross-talk differences exist in synaesthesia possible reflective of a broader or endophenotype. Interestingly, this broader phenotype appears to be confined to speech and thus language related processes, suggesting that language and synaesthesia may be closer connected, than the etymology of word ‘synaesthesia’ (perceive or feel together) implies.

Chapter 3. Synaesthesia and the McGurk Effect

Although reports of synaesthesia have ranged from tasting words (Ward & Simner, 2003; Ward & Simner, 2006) and sounds (Beeli, Essien & Jancke, 2005) to seeing calendar units (Smilek, Callejas, Dixon & Merikle, 2006) by far the most studied form is coloured letters, words or digits (Rich & Mattingley, 2002). This latter type of synaesthesia is diversely referred to as linguistic-colour synaesthesia (Simner *et al.*, 2006), lexical-colour synaesthesia (Rich *et al.*, 2005) or grapheme-colour synaesthesia (Ramachandran & Hubbard, 2001b) and such inconsistent terminology reflects an underlying lack of understanding about the amount of information processing required in order for synaesthesia to be induced.

Previous work in synaesthesia has provided evidence that synaesthesia is indeed a genuine perceptual phenomena (Kim *et al.*, 2006; Kim & Blake, 2005; Palmeri *et al.*, 2002; Ramachandran & Hubbard, 2001b; Smilek *et al.*, 2001) and not due simply to learned associations (Rich *et al.*, 2005), yet the amount of information processing of the inducer required to elicit a synaesthetic concurrent is unknown. Many studies have shown that synaesthesia can occur very rapidly (Dixon *et al.*, 2000; Palmeri *et al.*, 2002; Ramachandran & Hubbard, 2001b; Smilek *et al.*, 2001) and can be sensitive to changes in low-level image characteristics such as contrast (Hubbard *et al.*, 2006) or font (Witthoft & Winawer, 2006), suggesting that synaesthesia is a rapid, automatic association driven by sensory input. However, it has been demonstrated that attention (Mattingley *et al.*, 2001; Mattingley, Payne & Rich, 2006; Rich & Mattingley, 2003; Sagiv, Heer & Robertson, 2006) and semantic information (Dixon *et al.*, 2000; Jansari *et al.*, 2006; Myles *et al.* 2003, Smilek *et al.* 2006) play a role in mediating synaesthesia implying that a considerable amount of processing of the inducer is required to initiate the synaesthetic experience.

Related to these studies are cases where synaesthesia is thought to be an abnormal form of sensory integration (Grossenbacher & Lovelace, 2001; Marks, 1975; Ward *et al.*, 2006) or 'hyperbinding' (Hubbard, 2007; Mulvenna and Walsh, 2006; Robertson, 2003; Sagiv and Robertson, 2005; Weiss *et al.,* 2005). Among the evidence supporting the 'hyperbinding' hypothesis are two recent studies which both applied repetitive Transcranial Magnetic Stimulation (rTMS) on a number of synaesthetes. They found that when rTMS was applied to the right intraparietal sulcus of each synaesthete that this disrupted their synaesthesia as measured by the absence of the synaesthetic stroop effect (Esterman *et al.*, 2006; Muggleton *et al.*, 2007). Also, Weiss et al., (2005) found that IPS was differentially active in grapheme-colour synaesthetes, which they took to suggest that given that this region has been shown to be involved in multisensory integration that synaesthetes may have enhanced integration or "binding". These studies suggest that the processes underlying synaesthesia are situated more upstream than many of the models of synaesthesia suggest i.e. cross activation between cortical areas in the fusiform gyrus, and thus synaestheisa may be less a problem with low-level processing but may be dependent on more higher-level perceptual processes. To date, however, it is not known at what stage in information processing synaesthesia is induced and whether or not it is driven by the sensory input, or more higher-level perception output.

3.1 Experiment 4

In order to assess whether synaesthesia is triggered by early, sensory input or by relatively later perceptual processess, incongruent audiovisual recordings of spoken words were used, which were known to induce the 'McGurk' illusion (McGurk and McDonald, 1976). This illusion occurs when incongruent visual information (i.e. viseme) influences the perceived auditory information (spoken word), resulting in a change in the perceived sound from the real auditory input (Alsius *et al.*, 2005). Specifically, what was investigated was whether the colours induced by the spoken words are related to what is perceived (i.e. the illusory combination of audio and visual inputs) or to either of the individual sensory inputs (i.e. the viseme or phonemes) in 12 linguistic-colour synaesthetes. The reasoning was that if synaesthesia is driven by early sensory processing, then the colour induced to the audiovisual incongruent event would be related to the colour induced by either of the sensory modalities alone. On the other hand, if the synaesthetic colour is triggered relatively late in information processing, then the induced colour may not necessarily be related to the colours induced by either of the sensory components but could be a new colour.

3.1.1 Methods

3.1.1.1 Participants. The same synaesthetes who took part in Experiment 3 of Chapter 2 (section 2.3.1.1) took part in this experiment, which was carried out in the same session.

3.1.1.2 Stimuli and apparatus. The stimuli and apparatus were identical to Experiment 3 of the previous chapter. As mentioned in the previous chapter (section 2.3.1.4) a colour picker was used to enable the synaesthetes to pick a colour that best matched their synaesthetic colour. This colour picking program (programmed using Java) was presented on a different monitor. The program was mouse controlled and allowed participants to select from a wide array of colours defined by values along 3 (RGB) dimensions. Each chosen colour was recorded numerically as an RGB value (red, green, blue) ranging from 0 to 255, where the lowest value of 0, 0, 0 is white and largest value of 255, 255, 255 is black.

All stimuli were displayed using Presentation® software. Participants were tested individually in a windowless room with fluorescent lighting. They were seated in front of two monitors, one of the monitors (17inch) was positioned directly in front of each participant at a distance of 57cm with the result that the visual image of the AV display subtended a visual angle of 18 degrees. The experimental stimuli were presented through this monitor (1600*1200). The other monitor (21 inch) was positioned to the right of the participant and displayed the colour picker.

3.1.1.3 Design. The design was the same as in Experiment 3 of Chapter 2.

3.1.1.4 Procedure. The procedure was identical to the procedure of Section 2.3.1.4. Also the classification of responses was identical to Experiment 3 of Chapter 2 consisted of a fixation cross which was presented in centre of the screen for 1000ms. As the critical comparison in this experiment was whether the synaesthetic colour was different for each condition, a quantitative measure of colour difference was employed. This colour difference for each participant was measured as a vector difference and was calculated using the formula $\sqrt{}\{(R_1-R_2)^2 + (G_1-G_2)^2 + (B_1-B_2)^2\}$ (Ward *et al.*, 2006).

3.1.2 Results

Three synaesthetes reported no synaesthetic colour for any of the experimental stimuli, these synaesthetes had colours for graphemes but reported not to have any synaesthetic colours for words, either written or presented aurally. The remaining nine synaesthetes all reported synaesthetic colours in each condition.

For each particular word set (i.e. the set of three words comprising the audio word, visual word and the expected McGurk illusory word), three main patterns of results were found: for the A (audio) and incongruent AV (audiovisual) conditions participants either reported the audio for both (13% of responses), the illusory AV for both (17% of responses) or the correct audio for the A condition and the illusory AV response for the incongruent AV condition (43% of responses). To assess if the synaesthetic colour was determined by the multisensory percept or one of the unisensory inputs (for example the actual audio in the above cases) it was critical to examine the colour difference (using the RGB vector score, see 3.1.4) between the audio and audiovisual condition when the responses were appropriate to each condition.

Cases where the synaesthetes had the same colour for the critical phoneme of the audio and expected word of the audiovisual condition were excluded. For example for three synaesthetes NC, JF and AK the letter M and N had the same synaesthetic colour, and for many audiovisual trials (9 out of 30) the critical phoneme change was between a M and a N, see Table 1, consequently these cases were excluded from the synaesthetic colour analysis (the synaesthetic colour were each letter was determined in a separate study and that data was used to determine if a synaesthete had the same colour for two letters).

As a consistency measure for the reported colours the colours to trials where the audio word was reported at both the audio and audiovisual conditions were compared, (indicating that the McGurk illusion did not occur) and found that the mean colour difference across synaesthetes was quite low, mean RGB vector difference = 40.91. To check if this value merely reflected random variability in picking the same synaesthetic colour this colour difference value was compared to a consistency score obtained in a separate study to repeated letters. The resulting value was similar, mean RGB vector difference = 40.41, and they did not differ from each other, $t(16) = 0.07$, $p = .94$. Furthermore, when the illusory audiovisual word was reported for both the A and AV conditions, a similar result was obtained, mean RGB vector difference = 49.11. Again, this result did not differ from the consistency score, $t(16) = 1.16$, $p = 0.26$. Both results (40.41 and 49.11) also did not differ from each other, $t(16) = 1.01$, $p = 0.33$. These results are plotted on Figure 9.

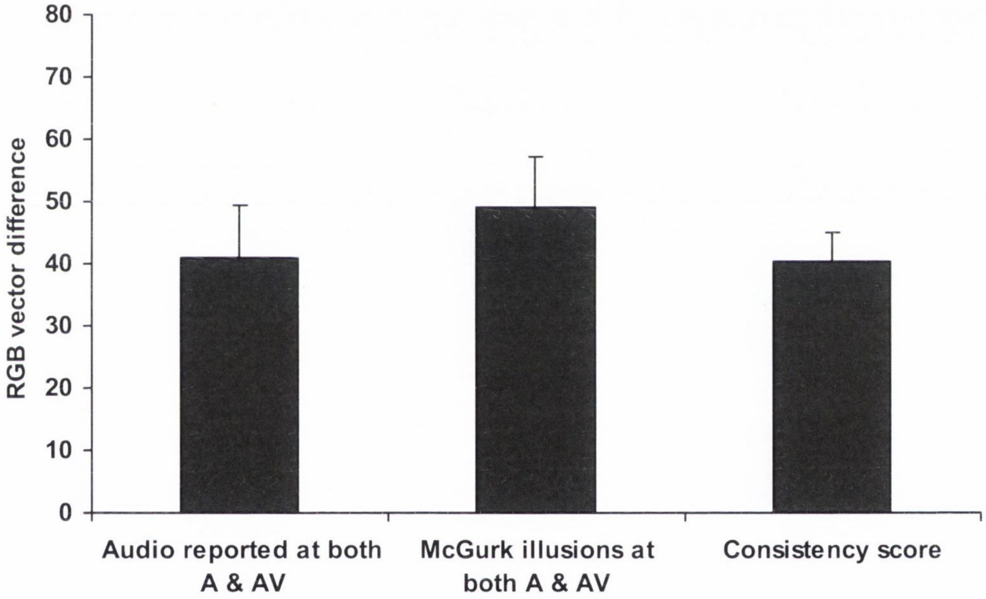

Figure 9: Plot illustrating RGB colour difference for three different response set comparisons (Experiment 4).
The synaesthetic colour comparisons (using a RGB vector score) were between; the colours reported when the audio was reported at both A & AV conditions (, e.g. stimulus: A = bait, AV = 'date'; response = 'bait' in both A & AV), and when the audiovisual word was reported at both A&AV conditions (, e.g. stimulus: A = bait, AV = 'date'; response = 'date' in both A & AV),. Also colour comparisons across a repeated picking of synaesthetic colours for letters were plotted (consistency score).

The RGB vector difference for the cases where the correct audio was reported in the A condition and the illusory audiovisual word at the AV condition was quite high, mean RGB vector difference = 212.27, and differed significantly from the consistency score for the repeated set of letters, $t(16) = 6.31$, $p < 0.001$. Excluded from this analysis were cases where the first letter didn't change between audio and audiovisual conditions upon McGurk illusions e.g. /been/ + [beep] = 'bea<u>m</u>'. In all these cases the colour didn't change given the strong effect of the first letter on the synaesthetic colour of the word.

As shown in Figure 10, the colour difference to the illusory AV responses relative to the correct A responses in the A condition, mean RGB vector difference = 212.27, was compared with the pooled difference when the same response (either the illusory audiovisual word was reported in both A and AV conditions or the correct audio was reported in both A and AV conditions) was given in both conditions, mean RGB vector difference = 47.51, which proved significant, $t(16) = 6.04$, $p < 0.001$.

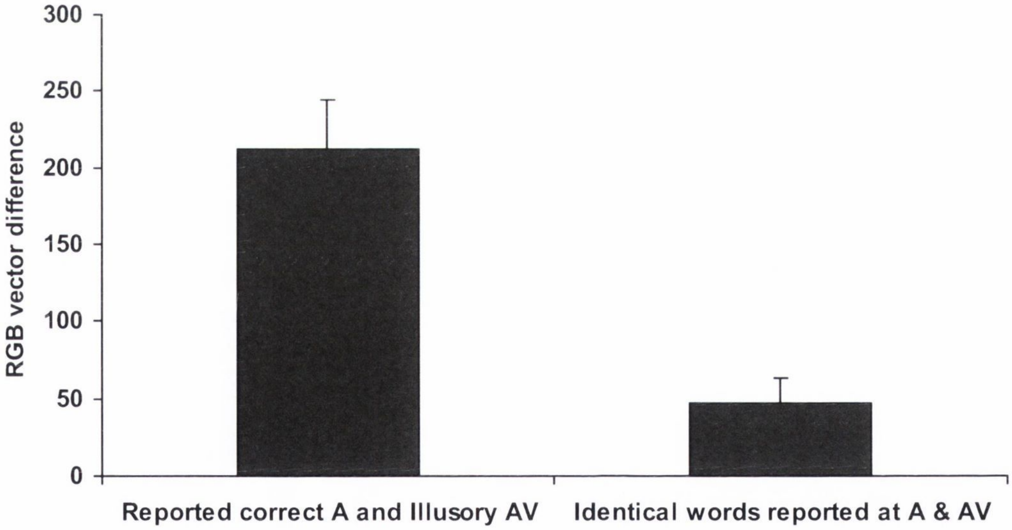

Figure 10: Plot illustrating RGB colour difference for two different response set comparisons (Experiment 4).
The synaesthetic colour comparisons (using a RGB vector score) were between; the colours reported when the correct audio and illusory audiovisual word were reported (reported correct A and illusory AV, e.g. stimulus: A = bait, AV = 'date'; response = 'bait' in A and date in AV)) and when the same word was reported for the audio and audiovisual conditions (identical words reported at A & AV, e.g. stimulus: A = bait, AV = 'date'; response = 'bait' in both cases)

It was possible that some synaesthetes might show a completely different pattern than the average pattern shown in Figure 10. An individual analysis on all the synaesthetes was carried out using the same comparison as in Figure 10 per synaesthete. This is shown in Figure 11.

All the synaesthetes, except one (MC), showed a change in their colour to the reported illusion relative to when the same response was given in both the AV and A conditions. MC had a very limited synaesthetic colour spectrum, which resulted in many instances where similar colours were reported to different (illusory and non-illusory) words, leaving too few data points for statistics. For the remaining 8 synaesthetes, the colour difference between the illusory audiovisual word in the AV condition and the correct audio in the A condition was significantly different from the colour difference when identical words were reported in the A and AV conditions (Figure 11).

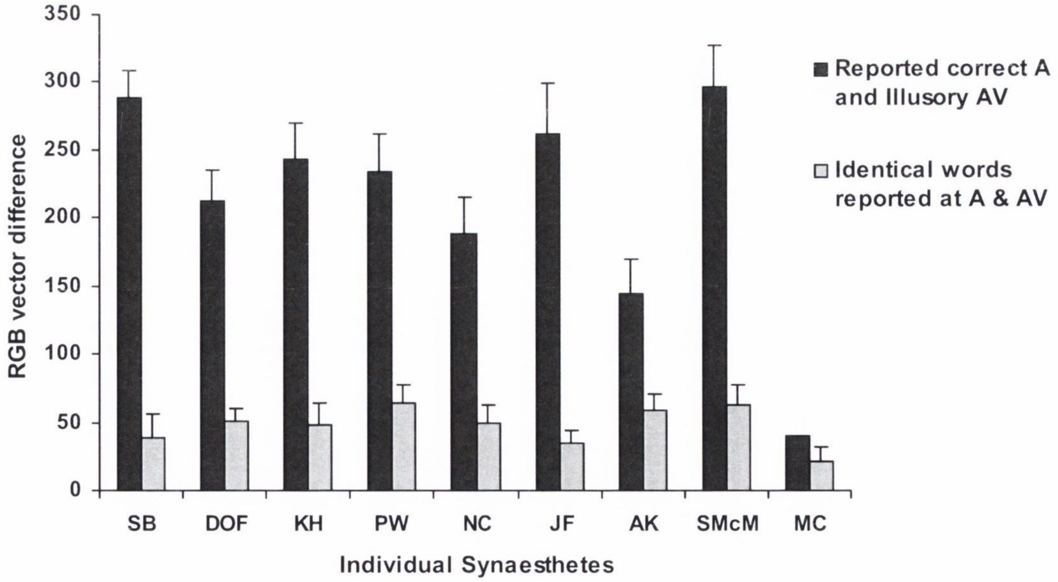

Figure 11: Plot illustrating RGB colour difference, per individual synaesthete, in RGB vector distance using the same comparison as in Figure 10 (Experiment 4).

In the V condition across all the trials for the 9 synaesthetes only 8 trials were correctly identified (i.e. 2.96%). There was also no case of a synaesthetic colour being reported in the V condition without a subsequent word being reported. It was not possible to compare the colours induced to correct V only trials with those induced to their AV counterparts because the correct response was often either the same as the viseme component or the fused percept in the AV trials. Of the remaining incorrect V only trials, either an unrelated word was reported or none was reported at all. When an unrelated word was reported the colour reported differed from the colour reported in AV condition, mean RGB vector difference = 191.10, which differed significantly from the consistency score of a repeated set of letters, $t(16) = 11.06$, $p < 0.001$.

3.1.3 Discussion

The critical question addressed was: upon undergoing the McGurk illusion would the synaesthetic colour match any of the unisensory inputs or would it match the multisensory percept? The main comparison was between the synaesthetic colour reported for the audiovisual word versus the synaesthetic colour of each of the unisensory inputs the audio and the visual. If the synaesthetes reported a different colour for the audiovisual condition than either of the unisensory conditions this indicated that the synaesthetic colour was dependent on the multisensory percept and not on the unisensory inputs. This is indeed the main pattern that was found.

As expected in the visual condition the synaesthetes could not identify the viseme word (2.96 % correct) and only reported a synaesthetic colour when a subsequent word was also reported. This suggests that the information in the V only trials was not sufficient either for correct identification of the word, or to induce a synaesthetic colour in the absence of identifying the word. In other words, this observation renders the idea that the visual component of the AV condition could be the sole contributor to the synaesthetic colour induced in the illusory AV trials as highly unlikely.

The most telling comparison involved the colour difference between AV and A conditions when the illusory audiovisual word was reported in the AV condition and when the correct audio was reported in the A condition compared with the colour difference between AV and A conditions when the same word was reported in both AV and A conditions (i.e. the pooled data: when either the illusory audiovisual word was reported in both A and AV conditions or the correct audio was reported in both A and AV conditions) (Figure 10). This result showed that the synaesthetic colour induced was in fact different than the colour induced for the audio when synaesthetes underwent the McGurk illusion.

An individual analysis for each synaesthete was carried out using the same comparison as above (Figure 11). This showed that for all but one synaesthete (this synaesthete had too few data points to do statistics) the same significant pattern was observed. Given the heterogeneity of synaesthesia (Dixon *et al.*, 2004; Dixon & Smilek, 2005; Hubbard, Arman *et al.*, 2005), there may have been a possibility that for some synaesthetes the synaesthetic colour would match the multisensory percept while for others it would match the actual unisensory audio. No synaesthete showed this opposite pattern of results.

Both critical comparisons show that for our linguistic-colour synaesthetes the colour induced by the McGurk illusion, i.e. the illusory audiovisual percept, was different than the colours induced to the unisensory components of the stimuli. Thus, the colour elicited to the McGurk illusion was not due to either of the unisensory inputs alone but was elicited by the integrated multisensory percept. This indicates that multisensory integration occurred prior to the manifestation of the synaesthetic association, i.e. visual information from the speakers face had already become integrated with the acoustic information before the synaesthetic concurrent occurred.

This finding may lend support to theories of synaesthesia which argue that synaesthesia occurs relatively late in the processing hierarchy either due to conceptual mediation (Dixon *et al.*, 2000), attention mechanisms (Mattingley *et al.*, 2001) or 'binding' mechanisms (Robertson, 2003) and consequently fits with neural models which implicate feedback as opposed to direct cross-activation between early sensory areas. As this study shows the synaesthesia occurs after the senses integrate or 'bind' the information, it follows that where the synaesthesia occurs in the processing stream is after the putative 'binding' step and thus does not involve early sensory areas.

This interpretation of the data, however, may be too simple. There is now a growing body of literature regarding the role of what were thought to be exclusively unisensory areas in multisensory integration (see Ghazanfar & Schroeder, 2006; Schroeder & Foxe, 2005; for recent reviews) and in audiovisual integration (Saint-Amour *et al.*, 2006). For example, in speech perception it has been shown through fMRI that the auditory cortex is active in silent lip reading (Calvert *et al.,* 2000) and more recently it has been shown conclusively that the primary auditory cortex is also active (MacSweeney *et al.,* 2000; Molholm & Foxe, 2005; Pekkola *et al.,* 2005). This activity is generally considered to be driven by feedback from a higher order multisensory region such as the STS (Calvert *et al.,* 1999), however evidence exists that this activity may be driven (at least partially) by feedforward activity from lower visual areas (Foxe *et al.,* 2000; Foxe & Schroeder, 2005; Schroeder & Foxe, 2002;). The function of auditory cortex activity in speech related processes is unclear, but it has been suggested that it could act as a priming mechanism to enable more efficient word recognition (Saint-Amour *et al.,* 2006) or simply improve speech recognition in noisy environments (Calvert *et al.,* 1999), whereas in the case of incongruent audiovisual information (as in the McGurk illusion) it could modify the acoustic information to resolve the audiovisual conflict (Saint-Amour *et al.,* 2006).

It is interesting to note the similarity in the nature of the arguments within audiovisual speech perception literature (and in multisensory integration in general) and the arguments within the synaesthesia literature i.e. whether the phenomena is the result of direct low-level sensory interactions or feedback from areas involved in higher cognitive functions, with rarely anything in between. For example with the McGurk illusion, given the automaticity and its resistance to cognitive interventions (Dekle, Fowler & Funnell, 1992; Fowler & Dekle, 1991; Green, Kuhl, Meltzoff & Stevens, 1991; Sams, Manninen, Surakka, Helin & Kättö, 1998), most researchers viewed it as case where integration automatically occurred early in sensory processing independent of higher-order processing. More recently, evidence has emerged that higher-order cognitive functions play a role such as attention (Alsius *et al.*, 2005) and sentence context (Windman, 2004). Windman reconciles what appear to be two opposing sets of data, by arguing that feedback from higher-order cortices can simply modify lower sensory areas and thus push the McGurk illusion in one direction or the other, but the processes involved in integrating/modulating the acoustic information occur at an early stage (given the strong evidence for primary cortex involvement in audiovisual speech integration).

Similarly within the synaesthesia literature the fact that conceptual information (Dixon *et al.*, 2000) and attention (Mattingley *et al.*, 2001; Mattingley *et al.*, 2006; Rich and Mattingley, 2003; Sagiv *et al.*, 2006) can modulate synaesthesia, can be similarly reconciled with the direct cross activation of early sensory areas by suggesting that feedback from areas involved in conceptual processing or attention modify the pattern of neural activation relating to the *synaesthetic inducer* and in this indirect way modify the concurrent synaesthetic experience. Hubbard and Ramachandran (Hubbard *et al*., 2006; Hubbard & Ramachandran, 2005, pg. 513-514) make a similar point discussing Dixon and others (Dixon *et al*., 2000) re-entrant model of synaesthesia which postulates feedback from conceptual areas in PIT (posterior intraparietal cortex) to the colour area (V4).

The involvement of parietal regions (right intraparietal sulcus) in synaesthesia as shown by two recent rTMS studies (Esterman *et al*., 2006; Muggleton *et al*., 2007) suggests that synaesthesia is a 'late' perceptual phenomenon. However, it may be the case that the involvement of parietal regions in synaesthesia may not be to play the role of the sole contributor to conscious 'binding' of extracted features as previously suggested (Hubbard, 2007), but instead may belong to a distributed net of connectivity involving lower sensory areas and possibly frontal regions which instead carries out a coordinated distributed exchange rather than 'binding' (which implies a temporal or spatial concordance) per se. In such a scenario the abnormal connectivity in synaesthesia could originate at an early stage in sensory processing but requires further downstream processing to manifest itself.

There is also another interpretation of the results of this study, which have not been discussed thus far: it is possible that once the synaesthetes integrate information (regardless of where or when this occurs) an extra step requiring grapheme re-encoding occurs and only then is the synaesthetic colour initiated. It has been shown that phonological processing can activate areas of the left posterior fusiform gyrus (Dietz, Jones, Gareau, Zeffiro & Eden, 2005; but c.f. Dehaence, Leclec'h, Poline, Lebihan & Cohen, 2002), very close to the visual word form area (VWFA) (Cohen &

Dehaene, 2004), so it is possible that graphemic representations could be elicited on presentation of spoken words. If the synaesthetic colour was linked directly to these graphemic representations then it could be the case that the integration of information (in the McGurk illusion) occurs before any synaesthesia related processes occur. Further studies testing explicitly the role of grapheme processing upon presentation of spoken grapheme or words need to be carried out to rule out or vindicate such an interpretation.

In summary, it was shown that in the case of the synaesthetes studied and where it was possible to get conclusive results that the synaesthetic concurrent matched the multisensory percept and was different from the unisensory inputs upon multisensory integration. This implies that multisensory integration occurred prior to the manifestation of the synaesthetic association and that the colour elicited to the McGurk illusion was not due to either of the unisensory inputs alone. These results lend support to 'late acting' theories of synaesthesia, where abnormal integration or 'binding' of information is implicated. However, caution most be exercised with such an interpretation, given our immature understanding of the cortical processes and cortical areas involved in the various forms of multisensory integration. Nevertheless, the data indicate that a significant amount of information processing is required before the synaesthetic experience is elicited and that synaesthesia is more likely associated with late perceptual processing (post-integration) rather than early sensory activation.

Chapter 4. The Role of Grapheme Processing in Linguistic-Colour Synaesthesia

Much debate in the synaesthesia literature has focused on the nature of the inducing stimulus, particularly the amount of information processing that is required to initiate a synaesthetic concurrent (Dixon *et al.*, 2000; Hubbard *et al.*, 2006; Ramachandran & Hubbard, 2001b; Rich & Mattingley, 2003). Different neuroanatomical models of synaesthesia have been proposed based on such assumptions regarding the critical components of the inducing stimulus. For example, in the cross-activation model of Ramachandran and Hubbard (2001a), it is argued that the visual grapheme processing is essential in linguistic-colour synaesthesia (at least with regard to what they call 'lower' synaesthetes (see Hubbard, Arman *et al.*, 2005)) and consequently propose that synaesthesia is the result of cross-activation from low-level grapheme areas to the adjacent colour area hV4. On the other hand, Dixon and others (2000) have found evidence that meaning or conceptual knowledge are important in linguistic-colour synaesthesia and thus suggest that higher order cortical areas (such as posterior infratemporal cortex), through re-entrant pathways, stimulate the colour area hV4 (Dixon *et al.*, 2000; Myles *et al.* 2003, Smilek *et al.* 2006).

In experiment 4, evidence was presented in support of synaesthesia as a late perceptual phenomena rather than one which is tied to low level properties of the inducing stimulus. However, a possible interpretation of the results of Experiment 4, an interpretation which would fit with the cross-activation of Hubbard and Ramachandran (2001a), would be that once audiovisual speech integration has occurred a further step involving grapheme conversion is required to initiate a synaesthetic colour. Thus any synaesthesia-related experience would take place after the audiovisual speech integration. It has been shown that phonological processing can activate areas of the left posterior fusiform gyrus (Dietz *et al.*, 2005) which is analtomically close to the visual word form area (VWFA) (Cohen & Dehaene, 2004). It is possible, therefore, that graphemic representations could be elicited on presentation of audiovisual speech stimuli. Indeed it has been shown that with auditory inducers (spoken words) that the synaesthetic colours are sensitive to graphemic information (Simner *et al.*, 2006). Simner and others (2006), using stress homographs such as '**con**-vict versus con-'**vict**, showed that the primary determinant of the synaesthetic colour was syllable stress, i.e. the primary vowel or the initial grapheme of the stressed syllable dictated the colour of the word. Syllable stress was shown to be critical for both spoken words and visually presented words. Interestingly, colour naming reaction times for words spoken were slower than their written counterparts, and the authors suggest this provides evidence that grapheme conversion is required for spoken inducers and the delay in colour naming is due to this necessary extra step for colour to be experienced. However, as the authors note, a methodological confound may exist given the intrinsic temporal processing differences (i.e. inter-sensory transduction rates) between acoustic and visual stimuli (Simner *et al.*, 2006).

Based on this evidence, Hubbard and others (2007) have recently proposed a theory, called 'the grand unified theory of synaesthesia'. This theory states that all forms of synaesthesia can be potentially explained by anatomically constrained cross-activation, where adjacent cortical areas cross-activate each other. This idea expands on Ramachandran and Hubbard's cross-activation theory (2001) which was proposed to explain linguistic-colour synaesthesia (or grapheme-colour synaesthesia). Given that phonetic areas of the brain and colour areas are not anatomically adjacent to each other, this theory states explicitly that with auditory inducers a second step requiring grapheme conversion occurs and only then is the colour initiated.

Many different types of synaesthesia are indeed compatible with the requirement that only adjacent cortical areas are involved in synaesthesia. Grapheme-colour synaesthesia is clearly compatible with the necessity of adjacency but less common varieties are also compatible. For example, as taste-shape synaesthesia is compatible as the somatosensory area for the tongue is adjacent to the taste area (Cerf-Ducastel, Van de Moortele, MacLeod, Le Bihan & Faurion, 2001) and lexical-gustatory synaesthesia is compatible since the phonetic and lexical areas lie close to the gustatory cortex in the insula (Ward *et al.*, 2005). Time-space synaesthesia could also be compatible with the adjacency requirement as it may involve adjacent parietal areas involved in numerical and spatial cognition (Hubbard, Piazza *et al.*, 2005). Likewise, ordinal linguistic personification has been suggested to be due to cross-activation between adjacent cortical areas in the inferior parietal lobule (Simner & Hubbard, 2006).

Interestingly, diffusion tensor imaging data on linguistic-colour synaesthesia has revealed increased connectivity (inferred from measured fractional anisotropic – how easily water diffuses along different axes) in the fusiform gyrus (as well as other cortical areas) providing evidence that cortical areas in the fusiform gyrus are exhuberantly cross-wired in synaesthetes (Rouw & Scholte, 2007). This would fit the idea that the critical cross-activation in synaesthesia is between a grapheme area and the colour hV4 since both functional areas reside in the fusiform gyrus. Also, given that there are structural connectivity differences in synaesthetes this implies, from a neurodevelopmental perspective, that one would expect cortically adjacent areas to be more affected in the amount of cross-connectivity than cortical areas further apart. As an important aspect of normal inter-cortical connectivity is that it is optimised to minimise wiring inefficiency (Cherniak, 1994; Cherniak, Mokhtarzada, Rodriguez-

Esteban, & Changizi, 2004; Chklovskii, Schikorski & Stevens, 2002; Klyachko & Stevens, 2003; Mitchison, 1992), thus neighboring cortical regions tending to be more heavily connected whereas distant cortical areas are less connected, making it much more likely that a genetic variant affecting cortical connectivity would impact on the connectivity between adjacent cortical areas. However, optimal wiring can be deviated when functional considerations are taken into account such as number of processing steps (Kaiser & Hilgetag, 2006).

If, in linguistic-colour synaesthesia, grapheme recoding of spoken inducers is not necessary to induce colours then this would suggest that auditory information is sufficient and the visual grapheme is not crucial in generating the synaesthetic experience. It would also provide evidence that the adjacency of cortical areas is not a requirement for all types of synaesthesia as the 'grand unified theory of synaesthesia' states. However, one observation suggests that a common link exists between grapheme and related phoneme inducers and that is that the same colour is invariably induced by both. This common link need not be the grapheme area but may be an area involved in more higher-level language related processing in the brain.

In the following study an interference paradigm was used to disrupt graphemic processing while synaesthetes listened to auditory inducers or observed visual inducers. The aim of the study was to assess whether grapheme processing is essential to colour generation in linguistic colour synaesthesia. Symbols which did not elicit a synaesthetic colour were chosen as stimuli which would interfere with grapheme processing. I first established in Experiment 5(a) and 5(b) if these symbols (i.e. Chinese characters) did indeed interfere with grapheme processing in a population of non-synaesthetes. In Experiments 6(a) and 6(b) these symbols were incorporated into an interference paradigm in order to disrupt grapheme processing when synaesthetes were presented with auditory (Experiment 6a) and visual inducers (Experiment 6b). A synaesthetic congruency effect was used in Experiment 6 (Mills *et al*., 1999; Odgaard *et al*., 1999) to assess the effect on the synaesthetic colours by the interference stimuli.

4.1 Experiment 5(a)

Before testing synaesthetes it was necessary to show that certain symbols interfere with grapheme processing. The symbols chosen were Chinese characters. These were chosen because of previous evidence in the literature that even non-native symbols can activate the same regions of the left fusiform cortex as do native graphemes (Callan, Callan, & Masakia, 2005). Importantly, Chinese characters did not elicit synaesthetic colours in any of our synaesthetes (as assessed prior this study). Upside-down letters were also considered as possible interference stimuli but many of the synaesthetes reported experiencing colour to these stimuli.

The interference stimuli were presented visually while participants were presented with auditory (and visual) inducers. To assess if such passive presentation of these visual stimuli actually affect auditory to grapheme conversion (or grapheme processing) a cross-modal matching task was employed, where participants had to match a spoken letter with its visual counterpart while an intervening visual stimulus (i.e. an interference stimuli) was presented between both. Performance in this matching task was assessed in terms of both speed (reaction time data) and accuracy (proportion correct). Upside down letters, face stimuli and visual white noise were all included with the Chinese characters as interfering visual stimuli. Upside down letters were included as a positive control as they shared all the features of normal graphemes and thus if they did not interfere with performance it would be indicative that the task employed was not sensitive enough. Visual white noise and face stimuli were also included as control stimuli i.e. as visual stimuli which are processed in different cortical regions than grapheme stimuli (Cohen *et al.*, 2002; Tarkiainen, Cornelissen & Salmelin, 2002; Puce, Allison, Asgari, Gore & McCarthy, 1996) and should not affect grapheme processing.

4.1.1 Methods

4.1.1.1 Participants. Eleven naïve volunteers (all female) participated in the experiment for pay. They ranged in age from 20 to 27 years, with a mean age of 22.18 years (SD = 2.32). All participants reported normal or corrected-to-normal vision and none reported any hearing impairments. Informed, written consent was obtained from all participants prior to the experiment.

4.1.1.2 Stimuli. The auditory stimuli consisted of 6 letter name sounds (i.e. A, H, G, E, R, T), recorded from the same actor, and each of which were modified (stretched or shortened) using audio software (Adobe Audition®) to be 400 msec long, without changing the pitch of the spoken. The visual stimuli consisted of the same letter stimuli and four sets of interference stimuli. The visual test stimuli were six lower case letters (a, h, g, e, r, t) each of which were presented for 100 milliseconds. Each letter subtended 7° by 4° to 7°of visual angle with the participant seated 58cm from the computer screen. The four different types of interference stimuli were; Chinese characters, upside down letters, face stimuli and visual white noise images, which were all also displayed for 100 ms each. There were six Chinese characters, six upside down letters, six face stimuli and one white noise image. The Chinese characters each subtended 8° by 4.5° to 8° of visual angle, whereas the upside down letters, each of which were different letters from the auditory letter names, subtended 7° by 4° to 7° of visual angle. The face stimuli were natural grayscale images of unfamiliar faces (3 female and 3 male faces), each subtending 8° by 9° of visual angle. The visual white noise image subtended 8.5° by 9° of visual angle. All the stimuli were presented on a 21 inch monitor (100Hz refresh rate). The experiment was programmed and the stimuli displayed using Neurobehavioural Systems Presentation® software.

A trial consisted of a fixation cross presented for a duration of 1500 msec in the centre of the computer monitor. This was followed by a 200 msec gap (i.e. where no stimulus was presented) and then the first auditory stimulus was presented (one of six letter names). After a stimulus onset asynchronies (SOA) of either 300 msec, 450 msec and 600 msec from the beginning of the auditory stimulus, a second stimulus, i.e. an visual interference stimuli from one of the four sets, was presented. Following the offset of this interference stimulus, the test stimulus was presented 50 msec after. The test stimulus comprised one of the six lower case letters. The test stimuli were either the same letter as the auditory letter or a different letter. The various SOA's were chosen to span a temporal window of 300msec where interference was considered most likely to occur, i.e. 300 msec into the 400 msec auditory stimulus, 50 msec after the auditory stimulus and 200 msec after the auditory stimulus.

4.1.1.3 Design. The experiment was based on a 3 by 4 within-subject design with SOA (300 msec, 450 msec and 600 msec) and interference stimuli (Chinese characters, upside letters, faces and white noise) as factors. The combination of factors gave a total of 12 different conditions.

Each one of these conditions contained 72 trials which included six different auditory letters repeated 12 times. Of these 12 repeated stimuli were matched to the visual test stimuli and half were different to the visual test stimuli. Trials across each of these 12 conditions were presented randomly (a total of 864 trials). A self-timed break was given every 100 trials. At the start of the experiment a block containing 30 practice trials were presented which contained trials randomly chosen from the main experiment. Data from practice trials were excluded from the analysis.

4.1.1.4 Procedure. Participants were tested individually in a windowless room with fluorescent lighting. A chinrest was used to ensure that the participants head faced straight ahead and that they maintained a distance of 57 cm away from the apparatus. Participants were instructed to indicate by key press, as quickly and as accurately as possible, whether the visual test letter was the same as the spoken letter heard or if it was different. The experiment took approximately 40 minutes to complete.

4.1.2 Results and Discussion

The dependent variables of interest were reaction times and accuracy (proportion correct) across both factors (SOA and interfering stimuli), errors and outliers (> 2.5 sd) were removed from the data analysis. See Figure 12 for the mean response times across SOA for all interference conditions.

A 3 by 4 within-subjects ANOVA was carried out on the reaction time data for all same and different trials. A significant main effect of SOA, $F(2, 20) = 21.644$, $p < 0.001$, was observed. The significant main effect of SOA reflected faster reactions times for the longer SOAs. This is most probably due to the temporal properties of the trial and the random design: with longer SOAs participants would have had more time to prepare a response and accordingly may have been quicker to press the response button. There was no significant main effect of condition, $F(3, 30) = 2.493$, $p =$

0.087, although it was approaching significance. However a significant interaction was found between the factors, $F(3, 60) = 2.363, p < 0.05$.

To investigate the interaction effect further, planned comparisons were conducted on the response times across each of the interference stimuli for each level of SOA. The 450 msec SOA appeared the most likely source of the interaction and planned comparisons confirmed that significant differences were observed between response times to the white noise stimuli and both Chinese characters, $F(1, 10) = 32.224, p < 0.001$, and upside letters, $F(1, 10) = 10.630, p < 0.01$. There was no significant difference found between response times to the white noise and face stimuli, $F(1, 10) = 2.013, p = 0.186$. This indicates that for the 450 msec SOA, Chinese characters and upside letters affect grapheme processing whereas neither white noise nor face stimuli (as expected) do not.

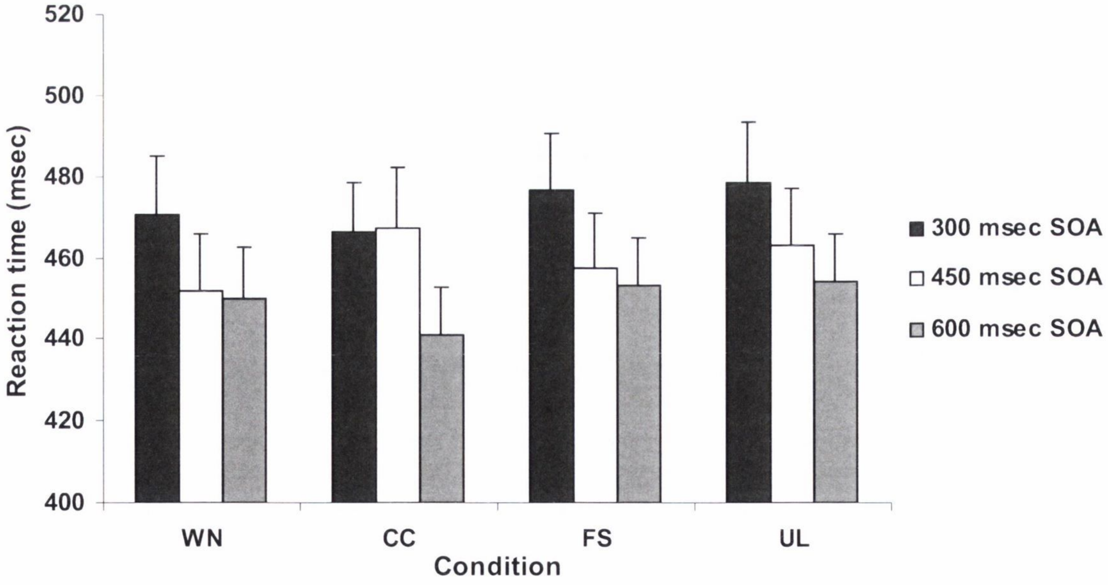

Figure 12: Plot illustrating reaction time data across each of the three SOAs (stimulus onset asynchronies) and the conditions (Experiment 5(a)).
WN = White noise, CC = Chinese characters , FS = Face stimuli and UL = Upside-down letters, collapsed over same and different trials.

A 3 by 4 within-subjects ANOVA was carried out on the accuracy data across both same and different trials which revealed no significant effects of SOA, $F(2, 20) = 2.768$, $p = 0.087$, or interference stimuli, $F(3, 30) = 1.542$, $p = 0.224$. There was no evidence of an interaction between the factors, $F(6, 60) = 0.561$, $p = 0.759$. However, when the same analysis was conducted on the 'same' trials only a significant main effect of interference stimuli, $F(3, 30) = 3.941$, $p < 0.05$ was found although there was no effect of SOA, $F(2, 20) = 1.194$, $p = 0.324$, nor an interaction, $F(3, 60) = 0.956$, $p = 0.463$.

Given one of the purposes of the experiment was to determine if in any of the SOA's Chinese characters and upside down letters were different from the other two interference stimuli planned comparisons in each SOA. This revealed that likely source of the main effect of the interference stimuli was based on a higher proportion of errors made to the Chinese characters and upside down letters conditions, in both the 450 msec SOA and 600 msec SOA trials, relative to the other interference stimuli (see Figure 13). This was particular true for the 450 msec SOA trials as Chinese characters and upside letters were both significantly different from the white noise condition, $F(1, 10) = 5.964$, $p < 0.05$, $F(1, 10) = 9.170$, $p < 0.05$, respectively, whereas face stimuli were not, $F(1, 10) = 0.904$, $p = 0.364$. Also, in the 600 msec SOA trials upside letters were significantly different than white noise, $F(1, 10) = 14.667$, $p < 0.01$, whereas Chinese characters versus white noise significance failed to reach significance, $F(1, 10) = 2.203$, $p = 0.169$. A possible explanation of why a difference between the conditions in accuracy was observed only in the 'same' trials was that participants responded more quickly to the 'same' trials versus the 'different' trials, $t(10) = 3.963$, $p < 0.001$, with a mean reaction time of 443 for 'same' trials and 477 for 'different' trials. This increase in speed lead to a drop in accuracy between the same and different trials, $t(10) = 4.245$, $p < 0.001$, with a mean accuracy score (proportion correct) of 0.938 for 'same' trials and 0.965 for 'different' trials. Most likely, this lower accuracy in the 'same' trials made differences between the interference conditions more readily observable, whereas in the 'different' trials, given the near ceiling level of performance, any differences between the conditions may not have manifested.

Both in reaction times and in accuracy rates ('same' trials only) the Chinese letters and upside down letters affected performance (slower reaction times and decreased accuracy) in the matching task relative to the other inteference stimuli, especially at a 450 msec SOA, and to lesser degree in the 600 msec SOA. This finding indicates that in grapheme processing (or grapheme recoding) these stimuli provide a level of disruption. Moreover, these findings suggest that processing of an auditory stimulus can be affected by a related visual stimulus (i.e. grapheme similarity) and that some level of information recoding may occur.

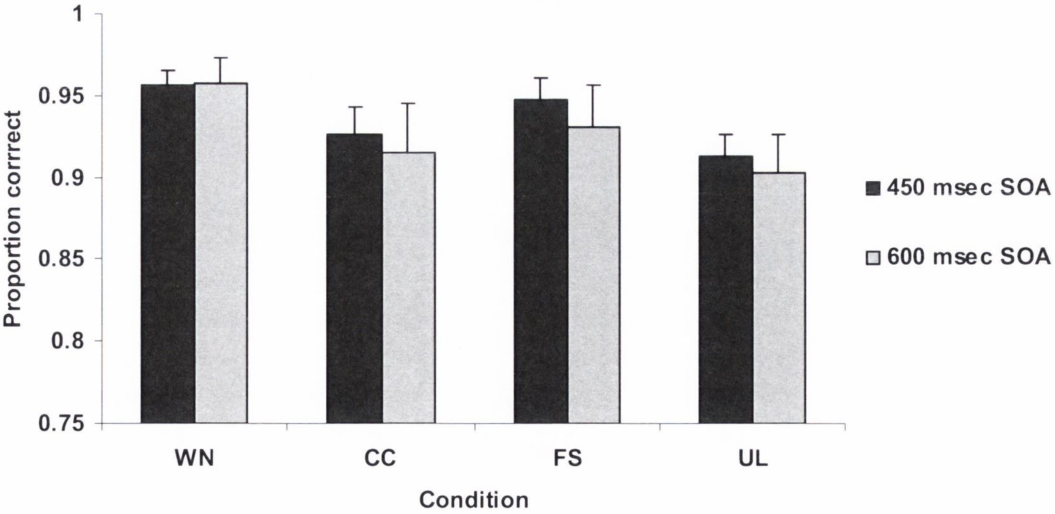

Figure 13: Plot illustrating proportion correct across each of the two SOA's (stimulus onset asynchronies) and each condition (Experiment 5(a)).

WN = White noise, CC = Chinese characters, FS = Face stimuli and UL = Upside-down letters, for same-only trials. The 300 msec SOA was excluded from this figure as the conditions showed little differences in performance.

4.2 Experiment 5(b)

Experiment 5(a) showed that Chinese characters disrupt auditory to grapheme conversion as assessed by a cross-modal matching task. Before testing synaesthetes to see if such interference with grapheme conversion affects the synaesthetic colour induced, it was necessary to carry out a similar matching task within the visual modality (i.e. presentation of visual graphemes rather than auditory inducers). Not only was it important to assess if interference holds up within the visual modality and for real grapheme presentation rather than auditory to grapheme recoding, but also because in synaesthetes both the auditory and visual modality were to be tested. The pattern of results obtained in the visual modality in synaesthetes would allow a comparison across modalities and thus determine whether a change in modality presentation of the inducers involves a different process in inducing the synaesthetic colour.

A similar matching task was used as in Experiment 5(a) to assess if the Chinese characters would interfere with grapheme processing, but here primary difference was that the first letters were presented within modality (i.e. visually) rather than acoustically.

4.2.1 Methods

4.2.1.1 Participants. Ten volunteers participated in the experiment for pay, (three male and seven female), none of which had participated in Experiment 5(a). They ranged in age from 22 to 40 years with a mean age of 27.1 years (SD = 5.28). All participants reported normal or corrected-to-normal vision and none reported any hearing impairments. Informed, written consent was obtained from all participants prior to the experiment.

4.2.1.2 Stimuli. The visual stimuli consisted of six uppercase letters (A, H, G, E, R, T), subtending 7.5° by 4° to 7.5° of visual angle. Each was presented for a duration of 100 msec. The interference and test stimuli (i.e. lower case letters) were identical to the stimuli used in Experiment 5(a). However, unlike the previous experiment, the interference stimuli were presented either 150 msec or 300 msec after the onset of the uppercase letters, for a duration of 100 msec. These SOA's were chosen because, with the first one i.e. the 150msec interference stimuli, this came on 50ms after the letter stimuli, which was as quickly as it could come on without disrupting the perception of the letter stimuli (while also being in line with the Experiment 5(a) SOA's), the 300msec SOA was chosen to remain consistent with Experiment 5(b). The test stimuli were presented 50 msec after the offset of the interference stimuli as in Experiment 5(a). All other aspects of the stimuli and the trials were identical to Experiment 5(a).

4.2.1.3 Design. The experiment was based on a 2 by 4 within-subject design with SOA (150 msec and 300 msec) and type of interference stimuli (i.e. Chinese characters, upside letters, faces and white noise) as factors. The combination of these factors gave a total of 8 different conditions. Each condition contained 72 trials; six different visual letters each presented six times with two types of test stimuli (i.e. same or different to the visual uppercase letters) giving 6 by 6 by 2 trials in each condition. Trials across the 8 conditions were presented randomly with a total of 576 trials in the experiment. Both the amount of trials between breaks and the amount of trials in the practice block were identical to Experiment 5(a).

4.2.1.4 Procedure. The procedure was identical to Experiment 5(a)

4.2.2 Results and Discussion

As in Experiment 5(a) the dependent variables of interest were reaction times and accuracy (proportion correct) across the two SOAs and four interference stimuli. Reaction times data across each SOA and condition for the same and different trials are shown in Figure 14. A 2 by 4 within-subjects ANOVA was carried out on the reaction time data which revealed a significant main effect of interference stimuli, $F(3, 27) = 12.087$, $p < .001$, no effect of SOA, $F(1, 9) = 3.50$, $p < 0.094$, and no interaction between the factors, $F(3, 27) = 1.512$, $p = 0.234$.

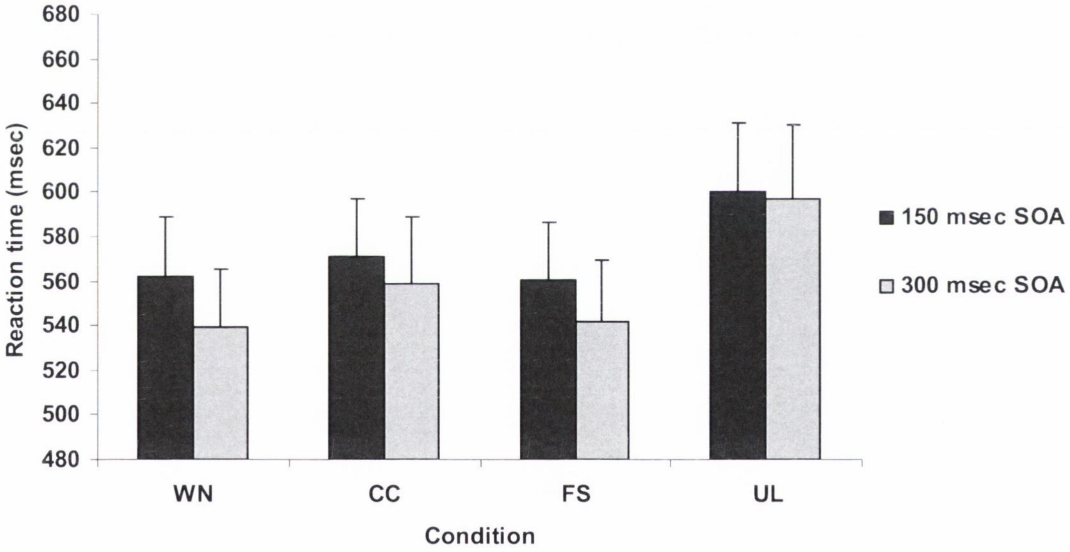

Figure 14: Plot illustrating reaction time data across each of the two SOA's (stimulus onset asynchronies) and the four conditions (Experiment 5(b)).
WN = White noise, CC = Chinese characters , FS = Face stimuli and UL = Upside-down letters, collapsed over same and different trials.

As in Experiment 5(a) in order to determine the SOA which showed the greatest difference in reactions times between the both the Chinese characters and the upside down letters versus the other interference stimuli (particularly the white noise stimuli), planned comparisons were carried out and revealed that the upside down letters and white noise stimuli were different, $F(1, 9) = 5.664$, $p < 0.05$, $F(1, 9) = 21.889$, $p < 0.01$, for the 150 msec SOA and 300 msec SOA respectively. Also the difference between Chinese letters and white noise may have contributed to main effect of condition although this difference failed to reach significance, $F(1, 9) = 4.566$, $p = 0.061$.

A 2 by 4 within-subjects ANOVA was carried out on the accuracy data collapsed across same and different trials, see Figure 15. Significant effects of SOA, $F(1, 9) = 10.154$, $p < 0.05$, and interference stimuli , $F(3, 27) = 10.768$, $p < 0.001$, were found but these factors did not interact, $F(3, 27) = 2.169$, $p = 0.115$. As in the reaction time data, planned comparisons revealed a significant difference between accuracy to the upside down letters versus white noise in both SOAs, $F(1, 9) = 13.554$, $p < 0.01$, $F(1, 9) = 5.165$, $p < 0.05$, for the 150 msec SOA and 300 msec SOA respectively. However, accuracy to the Chinese characters versus white noise was not different in either SOA, but in the 150 msec SOA it was the closest to significance, $F(1, 9) = 2.567$, $p < 0.144$.

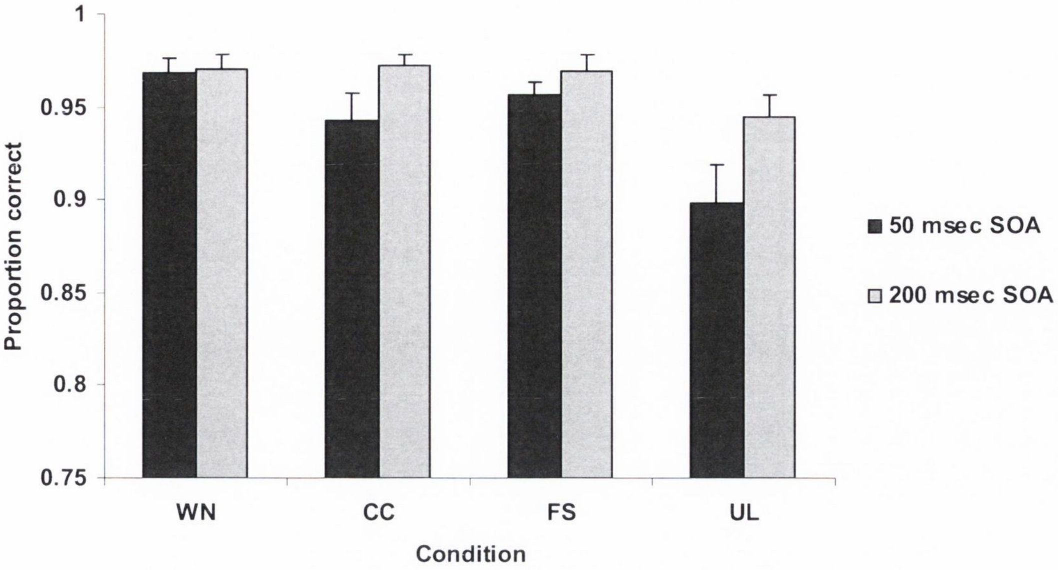

Figure 15: Plot illustrating proportion correct across each of the two SOA's (stimulus onset asynchronies) and the four conditions (Experiment 5(b)).
WN = White noise, CC = Chinese characters , FS = Face stimuli and UL = Upside-down letters, collapsed over same and different trials.

The results of this experiment suggest that in the visual modality, upside down letters significantly interfered with processing of other graphemes (in terms of both decreased accuracy and increased reaction times) relative to other interfering stimuli in the grapheme matching task. The Chinese letters also reduced performance but this effect failed to reach significance for both dependent variables.

4.3 Experiment 6(a)

It has been argued that in linguistic-colour synaesthesia auditory inducers induce a synaesthetic colour through an intermediate step which involves grapheme conversion of the auditory inducer (Simner *et al*., 2006; Hubbard *et al*., 2007). The aim of the following experiment was to test this idea by proposing that if this secondary process can be disrupted then this should affect induction of the synaesthetic colour. As was shown in Experiment 5(a), Chinese characters disrupt grapheme processing in a task requiring the matching of an auditory letter name to its visual counterpart (suggesting that the task involved conversion of an auditory stimulus to its graphemic form). Thus, it was expected that in synaesthetes visual presentation of Chinese characters upon presentation of auditory inducers should similarly disrupt any grapheme conversion process that is occurring and consequently affect the synaesthetic colour induced. Specifically, if processing of the inducing grapheme is disrupted then this may affect either the intensity of the synaesthetic colour or obliterate the synaesthetic colour entirely.

To measure the consequence of interference of auditory letter processing on the synaesthetic colour a synaesthetic congruency task was adopted (Dixon *et al.*, 2000; Mattingley *et al.*, 2001; Mills *et al.*, 1999; Odgaard *et al.*, 1999; Wollen and Ruggiero, 1983). This effect is based on a similar principle to the stroop task i.e. synaesthetes should be slower to name the colour of a grapheme presented in a colour incongruent to their induced synaesthetic colour compared with when the induced colour is congruent to the synaesthetic colour. This effect also works when the task is to name a colour patch that is either congruent or incongruent to the synaesthetic colour of an inducer previously presented (Dixon *et al.*, 2000; Mattingley *et al.*, 2006). Given that auditory inducers were involved this second type of paradigm was adopted.

The main prediction of the following experiment was that if recoding of an auditory letter to a visual grapheme is critical to synaesthetic colour induction then the magnitude of the synaesthetic congruency effect should be affected by an interfering visual stimulus. Specifically, a difference in the congruency effect was expected when Chinese characters were presented as an interference stimulus than white noise since in Experiment 5(a) Chinese characters were shown to interfere with auditory letter processing. An investigation of the differences in the magnitude of the synaesthetic congruency effect has been used before to assess the involvement of attention in synaesthesia (Mattingley *et al.*, 2006): a similar paradigm was adopted here.

4.3.1 Methods

4.3.1.1 Participants. Ten synaesthetes (10 female), ranging in age from 21-61 years, with a mean age of 37.3 years (S.D = 13.73) participated in this experiment for pay. All participants reported no hearing abnormalities and reported normal to corrected vision. All synaesthetes reported grapheme-colour synaesthesia, and have been previously tested for consistency in their grapheme to colour associations (see Barnett *et al.*, 2007). The study was approved by the School of Psychology Ethics Committee, Trinity College, Dublin. Informed, written consent was obtained from all participants prior to the experiment.

4.3.1.2 Stimuli and apparatus. The auditory stimuli used were identical to those used in Experiment 5(a) except that the auditory letters were chosen from a set of 10 in order to make sure the synaesthetes had different synaesthetic colours for each letter. The interference stimuli consisted of Chinese characters and a white noise image, both of which were identical to those described in Experiment 5(a). The test stimuli in this experiment consisted of colour patches subtended 8.5° by 9° of visual angle, which were created for each synaesthete individually using RGB values from colours picked using a colour picker.

The start of the trial, the duration of fixation and the presentation of the auditory letter stimulus, was identical to that described in Experiment 5(a). There were two types of interference stimuli (Chinese characters and white noise images) and they were presented at two different SOAs i.e. 450 msec SOA and a 600 msec

SOA. The test stimuli (colour patches) always came on 400ms after the offset of the auditory stimuli and this timing was kept constant to ensure that there were no differential delays between the auditory letter presentation and the test colour patch (as these delays may have affected the synaesthetic congruency effect). The colour patch remained on screen for four seconds or until a response was made. The test colour patch was either congruent or incongruent (an equal proportion of times) with the synaesthetic colour corresponding to the auditory letter, for each individual synaesthete. Catch trials consisting of animal pictures (dolphin, elephant, horse, rabbit, bird and dog), each subtending 8.0° by 8.5° of visual angle, were also presented infrequently in place of the interference stimuli.

A chinrest was used to ensure that the participant's head faced straight ahead and that they maintained a distance of 57 cm away from the monitor. A microphone (plantronics headset microphone) was attached to the chinrest to measure voice onset times. The threshold for the microphone (the sound pressure level which would trigger a response) was determined for each synaesthete.

4.3.1.3 Design. The experiment was based on a 2 by 3 within-subject design with SOA (450 msec and 600 msec) and type of interference stimuli (i.e. Chinese characters, white noise and no interference (no stimuli was presented)) as factors. This gave a total of six different conditions. Each condition contained 100 trials: six different auditory letters were each presented eight times with either a congruent or incongruent colour patch to the synaesthetic colour of the auditory letter. In addition four catch trials were included, giving a total of 100 trials. Trials across the six conditions were randomly presented giving a total of 600 trials in the experiment. Both the amount of trials between breaks and the amount of trials in the practice block were identical to Experiment 5(a).

4.3.1.4 Procedure. Participants were tested individually in a windowless room with fluorescent lighting. Participants were instructed to speak into the microphone as quickly and as accurately as possible the colour name corresponding to the test colour patch presented.

4.3.2 Results and Discussion

Microphone errors and outliers (> 2.5 sd) amounted to 3.1 % of trials, and thus were removed from the analysis. To assess for a synaesthetic congruency effect an analysis of the 'no interference' condition was carried out for both SOAs (450 msec and 600 msec) (see Figure 16). Slower reaction times for the incongruent condition versus the congruent condition would indicate a synaesthetic congruency effect. A 2 by 2 within subjects ANOVA was carried out on the data from the 'no interference' condition with both SOA and colour congruency (congruent and incongruent) as factors. No effects of SOA, $F(1, 9) = 0.538$, $p = 0.482$, or of congruency $F(1, 9) = 4.511$, $p = 0.063$, were observed. The main effect of congruency, however, was close to significance and analysis with planned comparison revealed that for the 600 SOA there was a significant effect of congruency, $t(9) = 2.286$, $p < 0.05$, whereas in the 450 SOA there was not $t(9) = 1.953$, $p = 0.083$. There was no interaction between SOA and congruency, $F(1, 9) = 2.651$, $p = 0.138$.

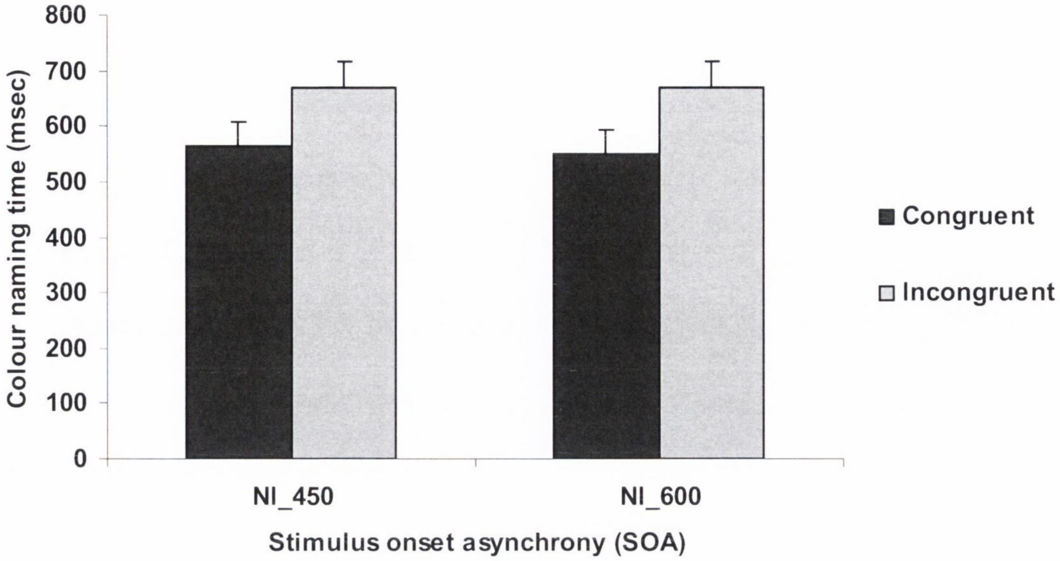

Figure 16: Plot illustrating mean colour naming times across each SOA for the 'no interference' condition (Experiment 6(a)).
NI_450 = No interference condition at 450 msec SOA. NI_600 = No interference condition at 600 msec SOA.

To determine if Chinese characters disrupted synaesthetic colour generation (by affecting grapheme processing), a 2 by 2 by 2 ANOVA was conducted, with SOA (450 msec and 600 msec), synaesthetic congruency (congruent, incongruent) and interference stimuli (Chinese characters, white noise) as factors, this is illustrated in Figure 17. This analysis revealed a significant effect of SOA, $F(1, 9) = 13.906$, $p < 0.01$, and no effect of either synaesthetic congruency, $F(1, 9) = 4.874$, $p = 0.055$, or interference, $F(1, 9) = 0.538$, $p = 0.482$. Also none of the four interactions were significant, with the critical two-way interaction of interest, that is the synaesthetic congruency by interference being, $F(1, 9) =0.020$, $p = 0.889$. The main effect of SOA was most likely due to temporal properties of the trial: although there was always a duration of 800msec between the onset of the auditory stimulus and the onset of the colour patch, there was a longer delay after the 450 msec SOA (the stimulus was presented for 100ms) until the colour patch onset (i.e. 250 msec) than in the 600 msec

SOA where there was only a gap of 100ms between the offset of the interference stimuli and the onset of the colour patch. This may have resulted in a longer preparation time in the 450 msec SOA to utter a response once the colour patch appeared, thus reducing colour naming times. This explanation is also in agreement with the lack of a main effect of SOA in the 'no interference' conditions.

Again the main effect of congruency failed to reach significance although it was approaching significance. Nevertheless, planned comparisons revealed that only one condition, the Chinese characters, at 600 SOA showed a significant synaesthetic congruency effect, $t(9) = 2.410$, $p < 0.05$. The other three comparisons did not reveal a significant synaesthetic congruency effect, $t(9) = 2.165$, $p = 0.059$, $t(9) = 2.221$, $p = 0.053$ and $t(9) = 1.946$, $p = 0.084$, for the Chinese characters and white noise at the 450 SOA and the white noise at the 600 SOA respectively. These comparisons reflect similar colour naming time differences between incongruent and congruent for the Chinese characters (M = 91.28 msec) and white noise (M = 111.33 msec) conditions at the 450 SOA, and for the Chinese characters (M = 116.04 msec) and white noise (M = 98.54 msec) conditions at the 600 SOA. Also, these mean differences are similar to those found for the 'no interference' condition in both 450 msec (M = 104.56 msec) and 600 msec (M = 118.85 msec) SOA's, indicating that there was no general influence of an intervening visual stimulus on the synaesthetic congruency effect.

The absence of the interaction between synaesthetic congruency and interference stimuli indicate that synaesthetes underwent the same amount of synaesthetic congruency regardless of the whether Chinese characters or white noise was the intervening stimulus. Thus the generation of the synaesthetic colour was unaffected by whether a visual stimulus was presented after the auditory inducer stimulus, even if this is a stimulus which can effect visual grapheme processing, However caution most be exercised in such an interpretation given that it depends on accepting the null hypothesis.

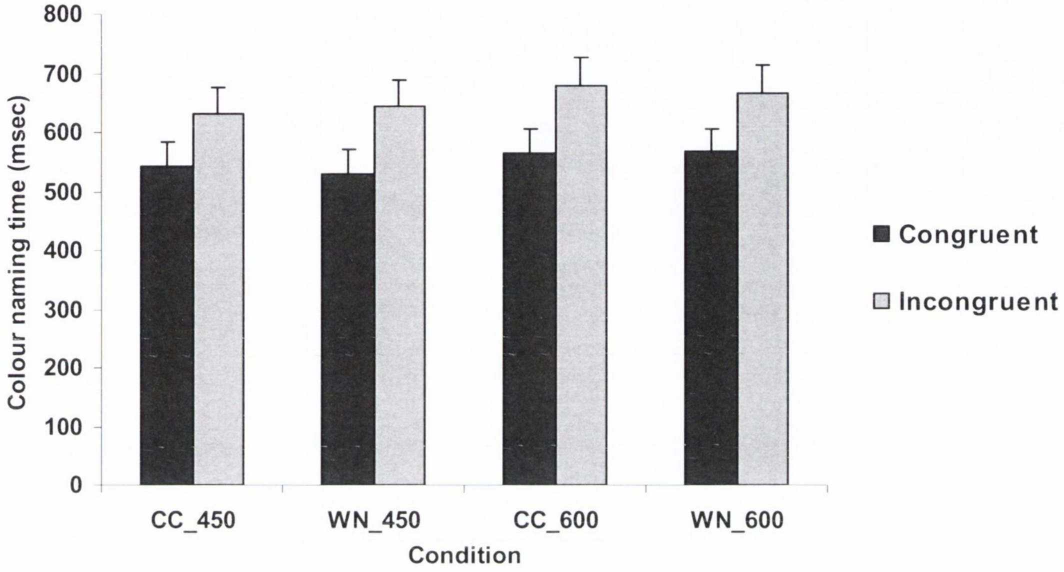

Figure 17: Plot illustrating the mean colour naming times across both SOAs and both interference conditions (Experiment 6(a)).
CC_450 = Chinese characters condition 450 msec SOA, WN_450 = White noise condition 450 msec SOA, CC_600 = Chinese characters condition 600 msec SOA and WN_600 = White noise condition 600 msec SOA.

4.4 Experiment 6(b)

The results of Experiment 6(a) provided evidence showing that when a visual stimulus (regardless if this stimulus affects grapheme processing or not) is presented during or after an auditory inducer, that this does not affect the magnitude of the synaesthetic congruency effect. Thus the processing involved between the auditory inducer and the synaesthetic colour induction is not likely affected by interfering visual information even if this visual information disrupts grapheme processing. Indeed the findings suggest that auditory induction of synaesthetic colours may not require an intermediate step involving recoding the auditory inducer into a visual or grapheme form but that the auditory inducer (or indeed a higher level linguistic abstraction of this inducer) may directly cross-activate a synaesthetic colour experience.

The following experiment was designed to investigate whether visual interference of the synaesthetic colour could occur when the inducer was also a visual stimulus. For example, if Chinese characters interfere in the visual modality and not the auditory modality with the synaesthetic colour induction (as measured by the synaesthetic congruency effect) this would suggest that visual graphemes have direct access to the induced colour but that there is a different process involved in auditory induction of colour. On the other hand, if the visual modality shows a similar pattern of results as in the auditory modality this might suggest that the synaesthetic colour is not tied to specific unimodal sensory information but is instead possibly tied to a more higher-level informational property of the letter category, for example a language-related property, which is independent of the encoding modality.

4.4.1 Methods

4.4.1.1 Participants. The same 10 synaesthetes who participated in Experiment 6(a) also participated in this experiment. The order of the experiments was counterbalanced across synaesthetes.

4.4.1.2 Stimuli. The visual letter stimuli used here were identical to Experiment 5(b) except that additional (to replace some of the six letters) letters were sometimes used to ensure that each synaesthete had different synaesthetic colours for each letter. The interference stimuli, test stimuli and catch trials were identical to Experiment 6(a), except that the interference stimuli were presented at an SOA of 150 msec from the visual letters. The fixation duration, breaks and practice blocks were all identical to Experiment 5(b).

4.4.1.3 Design. The experiment was based on a one-way, within-subject design with interference stimuli as the main factor (i.e. Chinese characters, white noise and no interference). For each of the three interference stimuli, each of the six different visual letters were presented eight times with either a congruent or incongruent colour patch, giving a total of 96 trials. Also, four catch trials were presented giving a total of 100 trials per each interference stimulus and a total of 300 trials across the experiment. Trials were randomly presented across participants and the amount of trials between breaks and the amount of trials in the practice block were identical to Experiment 5(a).

4.4.1.4 Procedure. The procedure was identical to Experiment 6(a) except that instead of an initial auditory inducing letter, here a visual letter was used. The task was the same as that described in the previous experiment.

4.4.2 Results and Discussion

Microphone errors and outliers (> 2.5 sd) amounted to 4.5 % of trials, and were removed from the analyses based on colour naming times. As in Experiment 6(a), the first step was to assess if the synaesthetes underwent a synaesthetic congruency effect in the 'no interference' condition based on the mean colour naming times. Although a significant difference was not found between the naming times to the congruent and incongruent trials, $t(9) = 2.176$, $p = 0.058$, it was approaching significance. The critical question, however, was whether a synaesthetic congruency effect was affected by the interference stimuli, Chinese characters and white noise. See Figure 18 for a Plot illustrating the response times across the congruency conditions to each of the interference stimuli. A 2 by 2 repeated measures ANOVA was conducted on the naming times with synaesthetic congruency (congruent, incongruent) and interference stimuli (Chinese characters and white noise) as factors. The effect of congruency approached significance, $F(1, 9) = 4.984$, $p = 0.052$, but there was no effect of interference stimuli, $F(2, 18) = 2.297$, $p = 0.129$, or no interaction between the factors, $F(2, 18) = 0.094$, $p = 0.910$. Critically, the failure to find an interaction between the factors indicates that there was no difference in magnitude of the synaesthetic congruency effect across the different types of interference stimuli. This is also evident in similar means obtained when the congruent condition reaction times are subtracted from the incongruent condition reactions times; Chinese characters condition (M = 114.11 msec), the white noise condition (M = 119.61 msec), and the ‘no interference condition (M = 116.1 msec).

As in Experiment 6(a) no difference in the synaesthetic congruency effect was found across the interfering stimulus types indicating that both types of interference stimuli had minimal effect on the synaesthetic colour generation and, more importantly, that they did not have a differential effect on the synaesthetic colour generation.

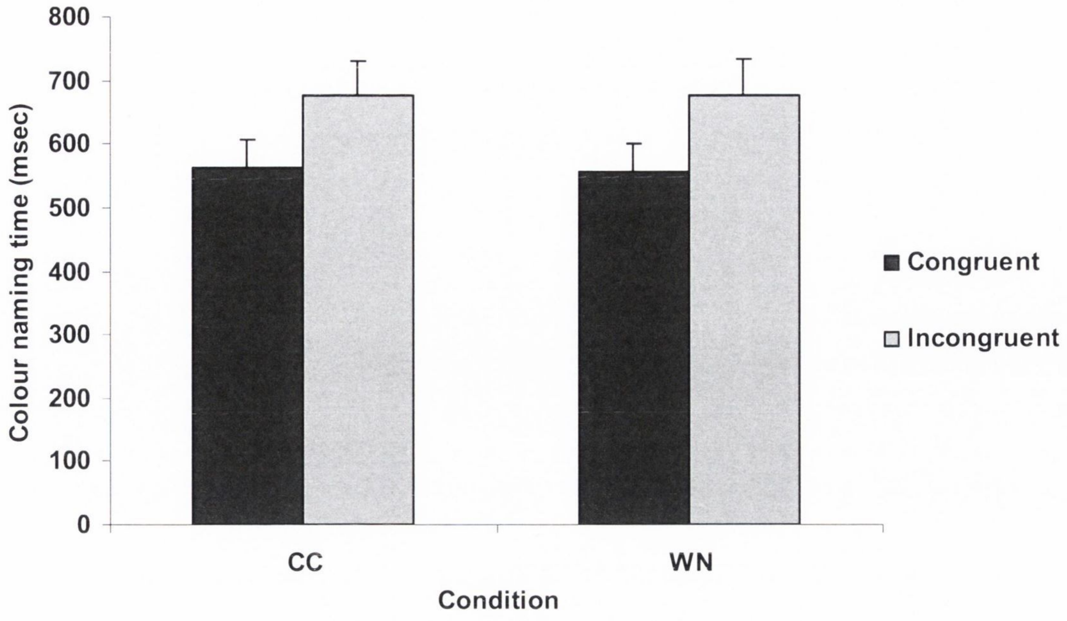

Figure 18: Plot illustrating mean colour naming times across both congruent and incongruent trials for each interference stimulus (Experiment 6(b)). CC = Chinese characters, WN =white noise.

As a further test to see if the grapheme has primary access to the synaesthetic colour, the degree of synaesthetic congruency (i.e. the difference in mean colour naming times between congruent and incongruent trials) was compared across this Experiments 6(a) and (b). If the grapheme is critical in linguistic-colour synaesthesia one may expect that a stronger congruency effect may occur for grapheme inducers (as in Experiment 6(b) here) rather than auditory inducers (Experiment 6(a)). Given that no interaction was observed between synaesthetic congruency and SOA, $F(1, 9) = 0.381$, $p = 0.552$, in Experiment 6(a) the mean colour naming difference between congruent and incongruent trials were averaged across SOAs for each type of interference stimulus. A 2 by 3 repeated measures ANOVA, with modality (auditory, visual) and interference stimuli (Chinese characters, white noise, no interference) as factors yielded no significant main effects of either modality, $F(1, 9) = 0.215$, $p = 0.654$, or interference stimulus, $F(2, 18) = 0.155$, $p = 0.857$. Also there was no interaction between modality and interference stimuli found, $F(2, 18) = 0.365$, $p =$

0.699. This analysis suggests that the degree of synaesthetic congruency is independent of the presentation modality of the synaesthetic inducer. Moreover, synaesthetic colours (as determined from the synaesthetic congruency effect) do not appear to be affected by the presentation modality, suggesting that higher level properties of the inducer (i.e. the linguistic associations of letters) stimulate the colour.

4.5 General Discussion

The primary aim of the experiments presented in this chapter was to investigate the role of grapheme processing in linguistic colour synaesthesia, particularly with regard to the induction of synaesthetic colours by auditory inducers (spoken letter names) and whether grapheme recoding of auditory stimuli is necessary in order to elicit the synaesthetic colour. To investigate this, an interference paradigm was adopted where symbols (i.e. Chinese characters) which did not induce synaesthetic colours were chosen as stimuli that may disrupt grapheme processing.

The first experiment (Experiment 5(a)) tested whether these symbols interfered with auditory letter processing using a cross-modal matching task in a group of non-synaesthetes. It was found (for same only trials) that performance was worse (slower reaction times and more errors) when the Chinese letters were presented as interfering stimuli relative to white noise as interference. Also performance to the Chinese characters stimuli was similar to upside down letters which acted as a positive control (some level of interference was expected with these stimuli). This provided evidence that Chinese characters were an effective stimulus to use to interfere with grapheme processing and that recoding of an auditory stimulus to its grapheme form likely occurred. Before testing synaesthetes a second matching experiment was carried, (Experiment 5(b)) investigating whether this effect also existed in within- modalities, i.e. when all stimuli in the task were presented visually, and found a similar pattern of results as those found in the cross-modal matching task. However, the comparison between the Chinese characters versus the white noise stimuli failed to reach significance in both accuracy and reaction times. This may have been down to the participant's use of a visual memory trace while carrying out the task which may have reduced the interference effect.

The next set of experiments (6 (a) and (b)) investigated whether the presence of Chinese characters would disrupt the generation of the synaesthetic colour, as measured by the synaesthetic congruency effect (Mattingley *et al.*, 2001). The congruency effect is found when synaesthetes are faster to name the colour of a colour patch if this colour is congruent to the synaesthetic colour induced by a preceding stimulus. The congruency effect was measured to letter stimuli presented in both the auditory (Experiment 6(a)) and visual modality (Experiment 6(b)) to assess if there would be differential effects across modalities that might, in turn, reflect different underlying neural processes for inducers presented in different modalities. In both experiments the critical interaction between synaesthetic congruency (congruent versus incongruent) and type of interfering stimulus was not observed, indicating that regardless of the nature of the visual interfering stimuli (Chinese characters or white noise), the same level of synaesthetic congruency was found. A comparison between the congruency effect found in Experiment 6(a) and Experiment 6(b) (i.e. the difference between colour naming times for congruent versus incongruent trials) also showed no difference. Also, the effect of synaesthetic congruency overall in both experiments did not reach significance (although it was marginal), which indicates a failure to replicate previous results (Dixon *et al.*, 2000; Mattingley *et al.*, 2001; Mattingley *et al.*, 2006; Mills *et al.*, 1999; Odgaard *et al.*, 1999). However, each condition showed virtually the same pattern of results, all of which were on the borderline of being significant. This borderline effect may have been due to the relatively few number of synaesthete participants in the task and, if more synaesthetes participates might have been significant.

The finding that congruency effects are not dependent on modality, and that visual stimuli do not interfere with the induced colour, suggest that recoding of auditory to visual grapheme is not a necessary step required to induce the synaesthetic colour (Hubbard *et al.*, 2007; Simner *et al.*, 2006). If this was the case then visual interfering stimuli should have disrupted this process and consequently affected the magnitude of the synaesthetic congruency effect. However it is possible that recoding does occur and that the interference stimuli did not affect this recoding process to a sufficient degree to hinder synaesthetic colour induction. This possibility cannot be ruled out, but the lack of a difference between modalities in the magnitude of the synaesthetic congruency effect suggests that regardless of the modality in which the inducing stimulus is presented, the synaesthetic colour affects performance in the same way. If the synaesthetic colour was dependent on the grapheme, as previous studies suggest (Hubbard *et al.*, 2006; Simner *et al.*, 2006), then it would be expected that the visual presentation of the grapheme (where no recoding is necessary) would cross-activate a colour more efficiently (i.e. faster and more accurately) than any process involving a cross-modal recoding of the stimulus. This was found not to be the case. Also, the presentation of a visual interfering stimulus did not affect the degree of a synaesthetic congruency effect compared with the absence of an interfering stimulus, again indicating that the synaesthetic colour initiation is not affected by low level sensory interventions.

The findings from the studies reported here are compatible with previous studies that show that the context of an ambiguous letter (Myles *et al.*, 2003) or the focus of attention (Mattingley *et al.*, 2006; Rich & Mattingley, 2003) can modulate or change a synaesthetic colour, or that the concept alone can initiate a synaesthetic colour (Dixon *et al.* 2000; Simner & Ward, 2006). All these studies suggest that the synaesthetic colour is not tied to the visual form of the grapheme but that what is critical is a higher level abstraction of the letter category (i.e. linked to more linguistic rather than visual properties). In order to accommodate such studies with the requirement of adjacent cortical areas, is that feedback (following attentional or conceptual modulation) from higher order areas could change the pattern of activation in the visual grapheme area itself thus affecting the synaesthetic colour produced (see Hubbard & Ramachandran, 2005). However, the results presented here also rule out this interpretation as it was shown that modulations of grapheme processing did not affect the synaesthetic colour as such an account would predict.

The experiments presented in this chapter have shown that if there is a cross-connection between the visual grapheme and the synaesthetic colour it is not tightly bound, since modulations of grapheme processing did not affect the synaesthetic colour.

Chapter 5. Examining 'Normal' Cross-Talk Between the Senses.

The previous chapters examined various aspects of synaesthesia, a condition commonly considered to be due to abnormal cross-talk between cortical areas. However, normal cross-talk between cortical areas is a widespread phenomena, one well studied example is multisensory integration. The aim of this chapter is to examine an aspect or offshoot of multisensory integration in order learn more about the processes involved in enabling or facilitating cross-talk between the senses with the hope that it can shed light on synaesthesia. For example, it has been shown that synaesthetes and non-synaesthetes share many similarities in the multisensory (and unisensory) associations they make (Sagiv & Ward, 2006; Smilek *et al.*, 2007; Ward *et al.*, 2006), suggesting that even if synaesthesia is due to an arbitrary variant in neural connectivity normal multisensory processes are playing some role (possibly working on top of such variant neural connectivity). Thus a complete understanding of synaesthesia may indeed require an understanding of how these multisensory processes enable such cross-talk.

Multisensory perception enables efficient processing of relevant information from across multiple senses (Ernst & Bülthoff, 2004), which is critical to successful behaviour. Each of the senses can deliver unique information regarding a certain perceptual attribute of an object or event in the environment and the task for the brain is to combine or integrate this information to enable accurate perceptual decisions. Sometimes, discrepancies may exist between these different sources of information across modalities often resulting in one of the sensory modalities dominating the percept.

For example, Welch and Warren, (1980, 1986) proposed the modality appropriateness hypothesis to explain the factors which affect multisensory perception. This hypothesis stated that, depending on the task, the most precise or accurate modality dominated perception. This sensory-dominance model has since been superseded by more probabilistic-based models where a mutual bias exists between sensory signals and the degree of dominance is dependent on the relative reliability information encoded by each sensory modality (Alais & Burr, 2004; Andersen, Tiippana, & Sams, 2005; Bresciani, Dammeier, & Ernst 2006; Ernst & Banks, 2002; Ernst, 2005; Knill & Pouget, 2004; Shams *et al.*, 2005). Thus, principles of sensory integration information are based mainly on the nature of the information itself (i.e. reliability) rather than the actual sensory encoding systems involved.

Certain multisensory illusions provide important insights into the mechanisms involved in multisensory integration, specifically with regard to the dominant information source for perceptual decisions. For example, when both an auditory and visual event occurs simultaneously but in different spatial locations, vision will often dominate the spatial percept (Colavita, 1979). Moreover, vision is said to 'capture' the location of the sound resulting in the auditory information being mislocated close to where the visual stimulus was presented (Bertelson & Radeau, 1981; Howard & Templeton, 1966; Jack & Thurlow, 1973; Welch & Warren, 1980). This phenomenon is known as the ventriloquist illusion and it is an example of a more general dominance of vision over other sensory modalities (Hay, Pick & Ikeda, 1965; Victor & Rock, 1964). However, in audiovisual tasks, vision only dominates (in accordance with probabilistic based models) when it is the more reliable signal but as vision becomes less reliable than audition (e.g. by increasingly blurring the visual image) then the variance of the visual signal is increased relative to the auditory input and audition then tends to dominate the spatial percept. This integration occurs in a weighted fashion where the relative weight allocated to each sensory system is inversely proportional to its variance (Alais & Burr, 2004).

Also the nature of the task is critical, in that, if it is of a spatial nature then vision is known to dominate perception (Rock & Victor, 1964) and if the task is of a temporal nature then audition is known to dominate both vision (Recanzone, 2003; Shams *et al.*, 2000; Shipley, 1964) and, to a lesser extent, haptics (Bresciani *et al.*, 2006). Ultimately, this dominance of spatial perception by vision and temporal perception by audition is down to the peculiarities of the sensory stimuli and the sensory systems involved (Witten & Knudsen, 2005).

Given that there is an inherent ability within each sense to process specific sensory information most efficiently, this raises the question of whether particular sensory information encoded by one modality is automatically processed by the most appropriate and reliable sensory modality in a mandatory manner. In other words, given that the auditory sense is most efficient at processing temporal information, we can ask whether temporal information would be automatically processed by audition irrespective of the original encoding modality? A recent study reported by Guttman, and others (2005) suggests that this is indeed the case. They found that when a sequence of visual stimuli (i.e. gabor patches) consisted of temporally random contrast changes participants tended to 'hear' the temporal order of these changes. This 'hearing' of temporal rhythms they referred to as a pseudo-synaesthesia in that visual temporal information *induces* hearing temporal rhythms. The reason for this as Guttman *et al.* (2005) argued was that since audition is the more appropriate or reliable modality for processing temporal information it then dominated the task, such that visual information was encoded into an auditory format. Using an interference paradigm they provided strong evidence for such mandatory cross-modal encoding of visual stimuli into the auditory domain.

In this study, whether the corollary also holds true was investigated: that obligatory cross-modal encoding occurs from audition to vision in tasks requiring spatial perception. In order to assess visual encoding of auditory spatial information a previously adopted interference paradigm used (Guttman *et al.*, 2005). Specifically, what was investigated was whether the perception of a sequence of auditory stimuli across a broad spatial array was affected by concomitant, task-irrelevant visuospatial information. The task involved comparing two successive spatial sequences of auditory events (i.e. whether the spatial order of white noise bursts was the same or different across two successive sequences). Given that what occurs during the encoding stage only of auditory spatial information was the primary interest, task-irrelevant visuo-spatial events were presented (i.e. LED lights) during the first and not during the second auditory sequence. In the visual sequence of events, the visual stimuli alternated between the auditory tones, resulting in visual sequences that were either congruent, incongruent, absent or non-spatial (i.e. on in all locations simultaneously) with the location of the auditory events. It was hypothesised that if vision is involved in encoding auditory spatial information then task-irrelevant visual information incongruent to the auditory spatial sequence should interfere with performance. In contrast, it was also expected that visual-spatial events that were congruent with the auditory events may facilitate performance. On the other hand, if spatial auditory information is most efficiently processed in audition, i.e. without the requirement of cross-modal encoding, then task-irrelevant visual information should have no effect on auditory spatial perception.

5.1 Experiment 7

It was important to first establish that the auditory spatial stimuli and the visual interference stimuli used in the subsequent experiments were perceptually distinct in both space and time, i.e. that visual capture or the 'ventriloquist effect' did not occur. It was necessary to ensure that any effects due to visual interference were not attributed to visual capture as this would render any differences in performance between the conditions as trivial. To eliminate visual capture, therefore, sufficient spatial and temporal disparities between each of the auditory and visual stimuli needed to be determined.

Spatial and temporal proximity is well known to facilitate the 'ventriloquist effect' (Jack & Thurlow, 1973; Radeau & Bertelson, 1977; Thomas, 1941; Thurlow & Jack, 1973). Indeed, according to Stein and Meredith (1993), temporal and spatial proximity make up two of the three principles of integration at the neuronal level (the third being the inverse effectiveness rule). Previous work has shown, for example, that introducing temporal disparities greater than 100ms between audio and visual stimuli results in a significant decrease in the 'ventriloquist effect' (Lewald *et al.*, 2001; Lewald & Guski, 2003; Radeau & Bertelson, 1987; Slutsky & Recanzone, 2001; Thomas, 1941). Moreover, the effect is virtually eliminated by temporal disparities of between 150-250ms (Lewald *et al.*, 2003; Slutsky & Recanzone, 2001). Other factors such as cognitive factors (e.g. when the participants were instructed to judge the likelihood of a common cause) can increase this temporal window of integration by up to 30 or 40 % (Lewald *et al.*, 2003), as can meaningful complex stimuli (speech and hand puppet movements for example) (Jack & Thurlow, 1973). However, the effect disappears at 300ms disparity even with complex stimuli (Jack & Thurlow, 1973).

Spatial disparitiy also has the effect of reducing the 'ventriloquist effect' especially for simple stimuli (Slutsky & Recanzone, 2001). Lewald and others (2001, 2003) found that a 3-4° separation between an auditory and visual stimulus was sufficient to detect that the two stimuli were perceptual distinct. On the other hand, Slutsky and Recanzone (2001) found that a 12° separation was sufficient to detect that the auditory and visual stimuli emanated from different locations, irrespective of temporal disparites (of e.g. 0, 50, 100, 150 and 250ms). However, these temporal and spatial disparities may nevertheless interact with each other. For example, although the audio and visual events may be perceived as being distinct, it is possible that for some temporal disparities the auditory percept may be perceived as shifted towards the location of the visual stimulus and away from its veridical location, but for larger temporal disparities this shifting effect should disappear (Slutsky & Recanzone, 2001).

White noise bursts were used as the auditory stimuli as white noise has been shown to be relatively easier to localise (Middlebrooks & Green, 1991). The assumption was made that by using auditory stimuli that were relatively easy to localise the visual stimuli would consequently have less of an interference effect on auditory localisation since we have a relatively high level of reliability in the auditory information for spatial localisation (Ernst & Banks, 2002): Slutsky and Recanzone (2001) reported that the more difficult it was to locate the auditory stimulus the more of an effect vision had on capturing the location of the auditory stimuli. As a further precaution to ensure that the visual stimuli did not capture the location of auditory events, each visual stimulus in the sequence alternated with the stimuli in the auditory sequence and each visual stimulus was presented after a temporal delay from the offset of each auditory stimulus. Moreover, a temporal delay between each auditory and visual event was used that was more than the temporal window of 50-100ms during which visual capture is known to occur see Lewald *et al.*, 2003; Slutsky & Recanzone, 2001) and which is thought to be related to transduction differences between the retina and the cochlea (Lewald *et al.*, 2003). To determine the most effective delay to avoid visual capture two different inter-stimulus delays of 250ms and 300ms were tested.

What was tested was whether the precautionary measures were successful in eliminating visual capture using a task where the participants had to compare the locations of a sequence of auditory events with the sequence of visual events which were embedded in the auditory sequence. If the participants could efficiently compare these spatial events across modalities then this would indicate that the visual and auditory stimuli were perceived as distinct and that visual capture was not occurring.

5.1.1 Methods

5.1.1.1 Participants. Ten psychology undergraduate students participated in the experiment for research credits (all were female). They ranged in age from 19 to 24 years (mean age = 21.2). All participants reported normal or corrected-to-normal vision and none reported any hearing abnormalities. All experiments reported here were approved by the School of Psychology Ethics Committee, Trinity College Dublin. Informed, written consent was obtained from all participants prior to the experiment.

5.1.1.2 Apparatus and stimuli. The apparatus consisted of six loud-speakers (4.5 cm Visaton FRWS 5), each mounted on wooden poles. The height of each speaker from the floor was 140 cm. A red LED was fixed on top of each speaker. An extra red LED was placed on top of a wooden pole (at the same height as the other

LEDs) and was positioned in the middle of the speaker array. This extra LED acted as a fixation point during the experiment. The speakers and LEDS were placed at fixed eccentricities from the location of the central fixation LED, with each stimulus subtending either 12, 36 or 60 degrees left and right of this central fixation. Each pole was positioned 100cm from the participant. A chinrest was used to ensure that the participant's head was facing straight ahead and that they maintained a distance of 100 cm away from each of the speakers in the apparatus. See Figure 19 for an illustration of the apparatus.

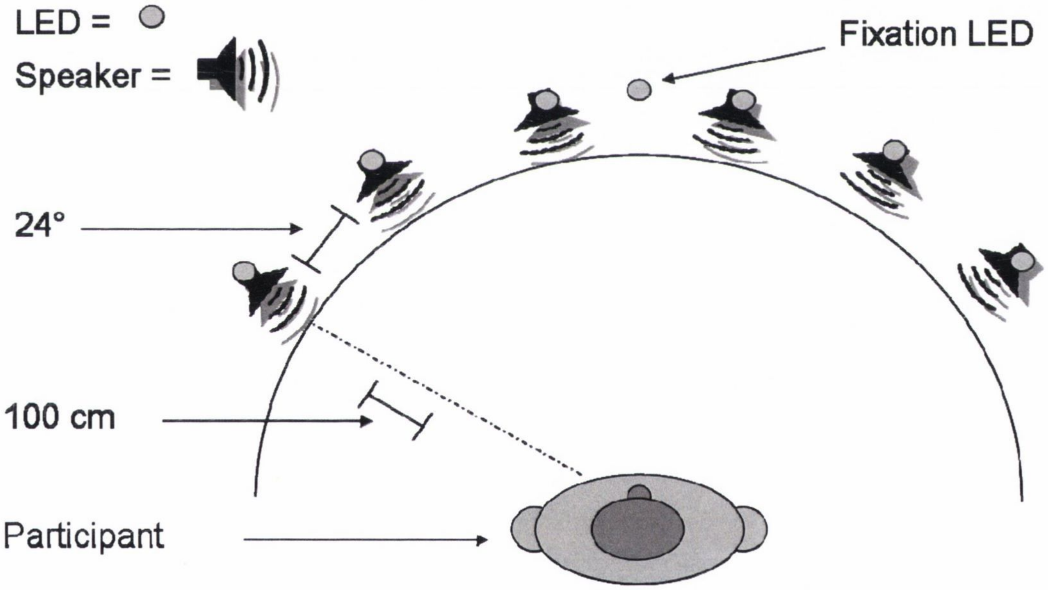

Figure 19: The apparatus used in all three experiments (Experiment 7-9).

The auditory stimulus was an 85dBA white noise burst (with a 2.5 ms ramp at either end of the sound wave envelop) with a presentation duration of 50 ms. A visual stimulus was a red LED also with a presentation duration of 50 ms. The auditory and visual stimuli were presented in a sequence consisting of six individual stimuli at different spatial locations. The presentations of each auditory and visual stimuli alternated with a temporal disparity between each audio and visual stimulus of either 250 ms in one set of trials or 300ms in the other. Each visual stimulus was presented after an auditory stimulus and was always presented at a time that was exactly mid-point between the offset of the previous auditory stimuli and the onset of the subsequent auditory event. Accordingly, the total delay between the offset of an auditory stimulus and the onset of the next auditory stimulus in the sequence in any one trial was either 600ms or 700ms depending on the onset of the visual stimulus (i.e. either 250ms or 300ms between each auditory event respectively). See Figure 20 for a schematic illustration of the temporal onset of the auditory and visual stimuli in a sequence of events.

An auditory spatial array consisted of a series of six noise bursts, each one following the other, in a sequence of random locations across the six speakers. The visual sequences were either spatially identical (i.e. each visual stimulus occurred in the same location as the preceding auditory stimulus) or they were presented in different spatial locations. When the auditory and visual stimuli occurred in different spatial locations, within the spatial array the first and last (i.e. the 6th location) auditory and visual pairs of the each sequence were in the same spatial location whereas the 4 middle auditory and visual stimuli were all in different spatial locations with a minimum spatial discrepancy of 48 degrees. Also, as a further constraint, each of the six spatial locations were used in creating the stimuli sequences. A small number of trials (10 %) acted as catch trials where the visual stimuli differed from the auditory stimuli in just two locations rather than four. These were introduced to eliminate effects due to simple response strategies, such as basing the response on only one or two of the audiovisual stimuli pairs in the sequence.

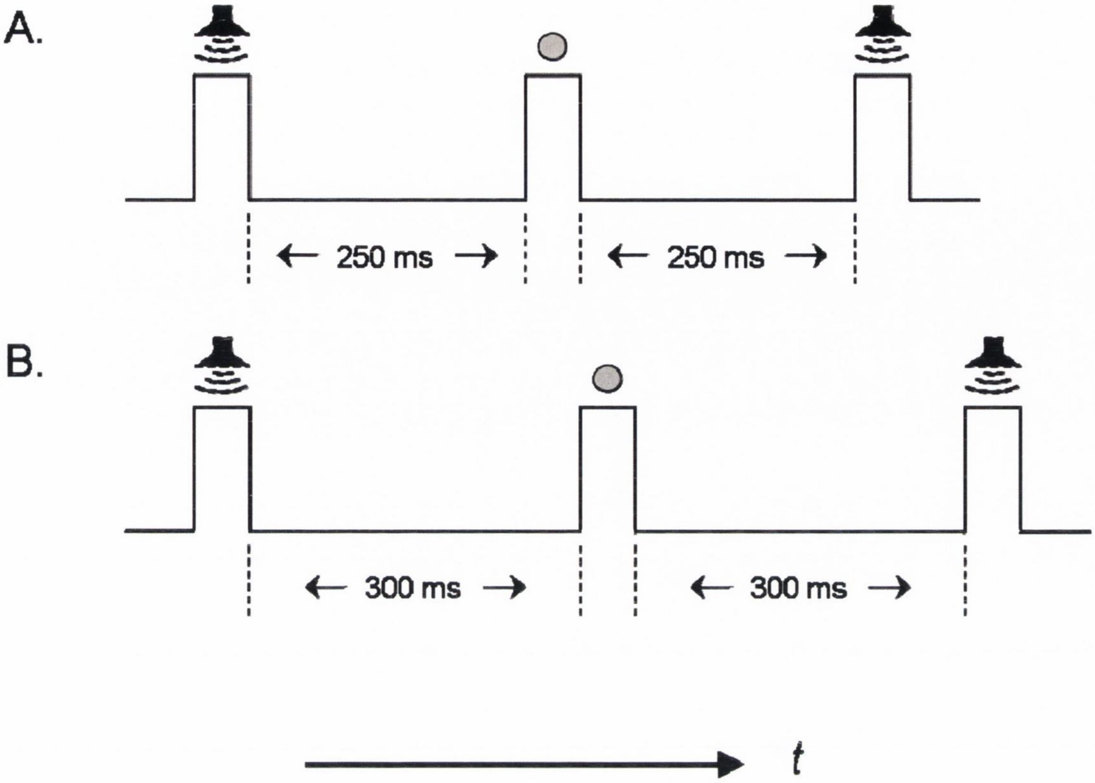

Figure 20: The temporal profile of the auditory and visual stimuli in Experiment 7.
(a) An Illustration of the temporal profile of the auditory and visual sequence when there was a 250 ms temporal disparity. The 300ms temporal disparity between AV events is illustrated in (b).

5.1.1.3 Design. The experiment was based on a one-way, within-subject design, with audio-visual temporal disparity as the main condition (either 250 ms or 300 ms).

The experiment contained 110 trials, 10 of which were catch trials and participants could take a self-timed break in the middle of the experiment. At the start of the each block participants were given 10 practice trials. The same auditory and visual sequences were used in both temporal disparities. For each temporal disparity there were 25 unique same trials, 25 unique different trials, and 5 catch trials. Trials were randomly presented across participants. Error rates were used (here and in subsequent experiments) as a measure of task performance.

5.1.1.4 Procedure. The experiment was conducted in a darkened room. Once the participant was guided to their seating position they were instructed to fixate on the central LED, (which appeared 1500ms before the experiment trials and remained on constantly) for the duration of the experiment. Participants were instructed to compare the spatial locations of auditory events with those of the visual events. In other words, their task was to decide whether the sequence of LED lights occurred in the same or different spatial locations of the white noise bursts heard. They were instructed to respond by key press and to respond as fast and as accurately as possible. Participants were also alerted that differences in the spatial order of stimuli across the auditory and visual arrays may often be based on one single location and not necessarily on all locations. Feedback was given in the form of a 50 ms, 250 Hz pure tone to incorrect responses only. A response triggered the onset of the next trial. The duration of a trial in each condition was 3350ms and 3900ms (for each of the 250ms and 300ms temporal disparities respectively) and the experiment took approximately 30 minutes to complete.

5.1.2 Results and Discussion

Figure 21 depicts performance (percentage correct) across the different temporal disparities collapsed over the same and different trials. The mean percentages for each disparity were 87.2% and 89.9% for the 250 ms and 300 ms respectively. Performance did not differ significantly across these disparities, $t(9) = 1.71$, $p = 0.121$, indicating that at either temporal disparity the participants can clearly distinguish between the auditory and visual sequence. In other words, since performance was relatively high for both disparities it gave an assurance that visual capture did not occur.

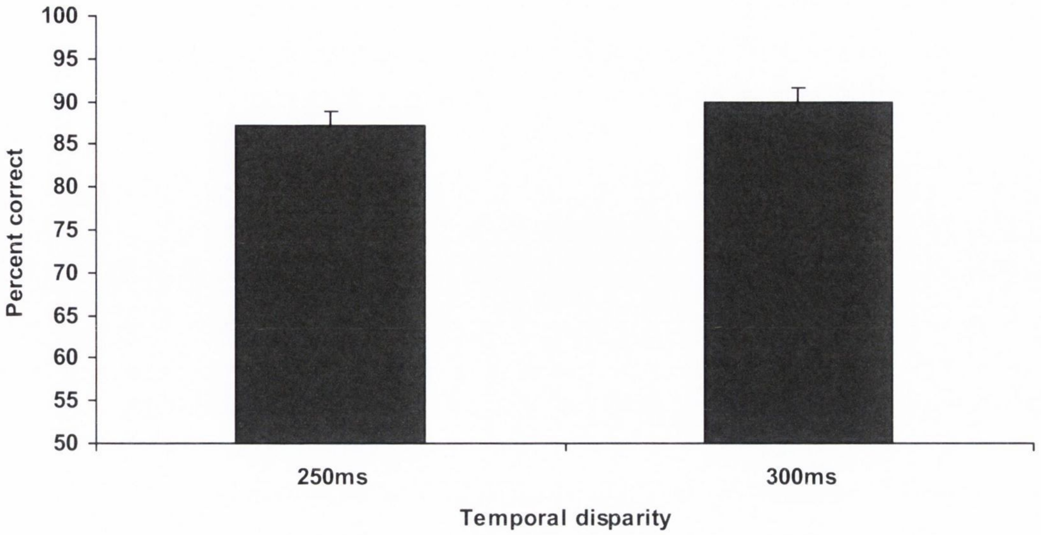

Figure 21: Plot illustrating the mean percentage correct performance in an audio-visual matching task across temporal disparities of 250 and 300 ms (Experiment 7).

Although not significant, performance at 300ms temporal disparity was marginally better than at the 250ms disparity. This probably reflects the overall temporal profile differences between the two trial types rather than any effect due to the ‘ventriloquist effect’. For example, the total duration of the trial when the temporal disparity was 250ms was shorter than for trials with a temporal disparity of 300ms. This meant that the participant had more time to judge any difference between the auditory and visual stimuli for the 300ms disparity. However, given the lack of a significant difference between both temporal disparities and that the results show no evidence of a ventriloquist effect, the temporal disparity of 250ms was used in the subsequent experiments.

5.2 Experiment 8

To assess whether auditory spatial perception involves visual encoding an experiment was designed where task-irrelevant visuo-spatial information was embedded into an auditory spatial matching task and investigated whether this visual information interfered with auditory spatial processing. The main hypothesis was that if the spatial discrimination of auditory sequences involves visual encoding then incongruent visual information, even if it is irrelevant to the task, should disrupt performance relative to a condition where no visual information is presented (i.e. auditory-only condition). Moreover, it was expected that congruent task-irrelevant visual information might facilitate performance relative to the auditory-only condition. For comparison, a non-spatial visual condition where all visual stimuli were presented simultaneously across all locations during the auditory events, was included. This condition allowed an assessment of whether visuo-spatial information specifically or visual information generally interferes with audio-spatial processing.

A trial consisted of a spatial sequence of locations defined by auditory noise bursts followed by a second such auditory sequence which was either identical or different to the first. The participant's task was to decide if these pairs of sequences were the same or not. During the first auditory spatial sequence an auditory stimuli was interleaved with task-irrelevant visual stimuli (i.e. LEDs at various spatial locations). These visual sequences were presented during the first of the auditory sequence as the main interest was the level of interference at the encoding stage. A previous related study on audio-visual interference in the temporal domain found that it was at the level of encoding rather than at the retrieval stage where the main cross-modal interference occurred (Guttman *et al.*, 2005).

5.2.1 Methods

5.2.1.1 Participants. Twelve psychology undergraduate students participated in the experiment for research credits (one male, 11 female). They ranged in age from 19 to 21 years (mean age = 19.67). All participants reported normal or corrected-to-normal vision and none reported any hearing impairment. Informed, written consent was obtained from all participants prior to the experiment. None had participated in the previous experiment.

5.2.1.2 Apparatus and stimuli. The apparatus and individual auditory and visual stimuli used were the same as those described in Experiment 7.

5.2.1.3 Design. The experiment was based on a one-way, within-subject design, with auditory-visual congruency as the main factor. There were four levels to this factor: congruent, incongruent, non-spatial vision and no visual information. Trials in the congruent and incongruent conditions were randomly presented in the same block whereas the non-spatial vision and auditory-only trials were presented in separate blocks. These blocks were counterbalanced across participants. Each block contained 110 trials (10 of which were catch trials). At the start of the experiment participants were given 10, randomly presented practice trials in each of the conditions.

5.2.1.4 Procedure. The task for the participant in each trial was to compare two successive auditory spatial sequences (each consisting of six individual stimuli). These sequences were separated by a delay of 1800 ms. In the first of the two auditory sequences, each of the six visual stimuli were presented alternating between the auditory stimuli with a temporal delay of 250 ms from the offset of each auditory stimulus (as determined in Experiment 7). Figure 20(a) provides a schematic illustration of the temporal onset of each visual stimulus between any two auditory stimuli in the auditory sequence.

According to each condition, the visual stimuli either occurred in the same location as each of the auditory stimuli in a sequence (i.e. congruent), or, as described in Experiment 7, four of the six of the visual stimuli occurred in a location other than that of its auditory counterpart (i.e. incongruent). In the non-spatial visual condition all visual stimuli in the array were presented simultaneously between each auditory event. Finally, the no-visual stimulus condition (i.e. auditory only) served as the baseline performance indicator in this task.

Within a trial, therefore, visual stimuli occurred within the first auditory sequence but not in the second. The spatial order of stimuli in the second auditory sequence was either identical or different to the spatial order of events in the first auditory sequence. When the auditory sequences were different the first and last stimuli of the sequence were the same whereas in the remaining locations they differed with a minimum spatial discrepancy of 48 degrees. Participants were instructed to ignore the visual information and compare the order of spatial locations across the two auditory spatial sequences within a trial. There were equal numbers of same trials and different trials in each condition (i.e. 25 trials each). As in Experiment 7, catch trials were introduced where the first and second auditory sequence differed only in two of the six locations in a sequence rather than four, to eliminate simple response strategies. Feedback was also presented after each incorrect response. The experiment took approximately 1 hour and 40 minutes to complete (carried out over two sessions).

5.2.2 Results and Discussion

The mean percentage correct responses across the four conditions; congruent, incongruent, non-spatial and auditory-only were 79.25%, 71.42%, 71.67% and 81.5% respectively. Performance is illustrated in Figure 22. A repeated measures ANOVA was conducted and revealed a significant difference in performance across conditions, $F(3, 33) = 7.58$, $p < 0.01$. A post-hoc, Newman-Keuls test revealed that performance in the congruent, non-spatial and auditory-only conditions did not differ from each other. However, performance in the incongruent condition was significantly worse than that in the congruent, $p < 0.01$, non-spatial, $p < 0.01$ and auditory-only conditions, $p < 0.001$. This finding indicated that visual information incongruent to the auditory sequence significantly interfered with performance in the auditory domain. Moreover, this effect cannot be attributed solely to the presence of visual stimuli since the non-spatial visual and congruent condition both contained visual information and no interference effect was observed for these conditions.

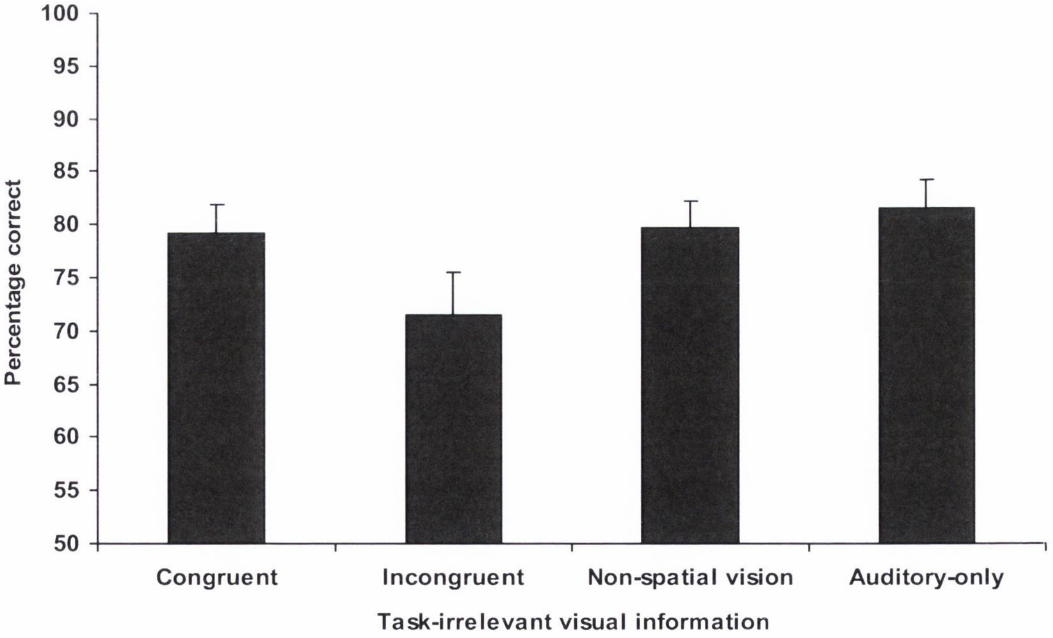

Figure 22: Plot illustrating the mean percentage correct performance across the four different conditions (Experiment 8).
Namely, congruent, incongruent, non-spatial visual and no visual information only.

Interestingly, however, a facilitation by the congruent condition was not observed as might be expected if auditory spatial information is encoded into a visuo-spatial format. Because of the relative reliability of visual events for spatial localisation, it was expected that congruent visual information should facilitate localisation of auditory events in the same way in which incongruent visual information impedes it. It was suspected that the lack of a facilitation in the congruent condition may have been due to the block design employed (i.e. trials in the non-spatial visual and no visual information conditions were blocked whereas trials in the congruent and incongruent conditions were presented in the same block.) This may have resulted in a change in the response criteria adopted across blocks and a more cautious approach may have been required in the mixed congruent and incongruent block due to increased uncertainty, thus preventing a facilitation from a occurring to the congruent trials. To test this idea the following experiment was designed in which trials in each of the congruent and incongruent conditions were presented in separate blocks.

5.3 Experiment 9

In this experiment trials were presented in separate blocks according to each condition. Here it was predicted that a facilitation would occur for the congruent condition compared with the non-spatial visual condition. Also expected was a relative cost in performance in the incongruent condition as was found in Experiment 8. In contrast to the previous experiment an auditory-only condition was not included and the non-spatial visual condition was as a control condition (performance was almost identical across both of these conditions in Experiment 8). Instead, the non-spatial visual condition provided a measure of the necessity of spatial information in the visual stimuli in affecting auditory processing.

5.3.1 Methods

5.3.1.1 Participants. Twenty one undergraduate students participated in the experiment for research credits (4 male, 17 female). They ranged in age from 19 to 46 years (mean age = 22.38). All participants reported no hearing impairments and also reported normal or corrected-to-normal vision. Informed, written consent was obtained from all participants prior to the experiment. None had participated in the previous experiments.

5.3.1.2 Apparatus and stimuli. The apparatus and stimuli were identical to Experiment 8 except that, for practical reasons, the auditory stimuli were presented at 80dBA sound pressure level (some participants in a pilot study reported feeling uncomfortable with the original decibel level of the audio sounds).

5.3.1.3 Design. The experiment was based on a one-way, within-subject design, with audio-visual congruency as the main factor (with 3 levels of congruent, incongruent, and non-spatial vision). Trials in each condition were presented in a different block and block order was counterbalanced across participants. Each block contained 55 trials comprising of 25 unique same trials, 25 unique different trials and 5 catch trials. At the start of the experiment participants were given 15 practice trials (5 trials from each of the 3 conditions) presented in a random order across participants.

5.3.1.4 Procedure. The procedure was identical to that described in Experiment 8. The participants were not made aware that each block contained similar trials.

5.3.2 Results and Discussion

Performance across the congruent, incongruent and non-spatial visual blocks is illustrated in Figure 23. The mean percentage correct performance was 77.14, 67.14, and 72.57 for the congruent, incongruent and non-spatial conditions respectively. A one-way, repeated measures ANOVA, $F(40) = 6.88$, $p < 0.01$, indicated a main effect of condition. Planned comparisons revealed that performance in the congruent condition was significantly better than performance in either the incongruent, $F(20) = 9.08$, $p < 0.01$] and non-spatial, $F(20) = 4.51$, $p < 0.05$, conditions, suggesting that performance was facilitated during the congruent condition. As in Experiment 8, a cost was observed for the incongruent condition in comparison to the non-spatial visual condition, $F(20) = 4.76$, $p < 0.05$.

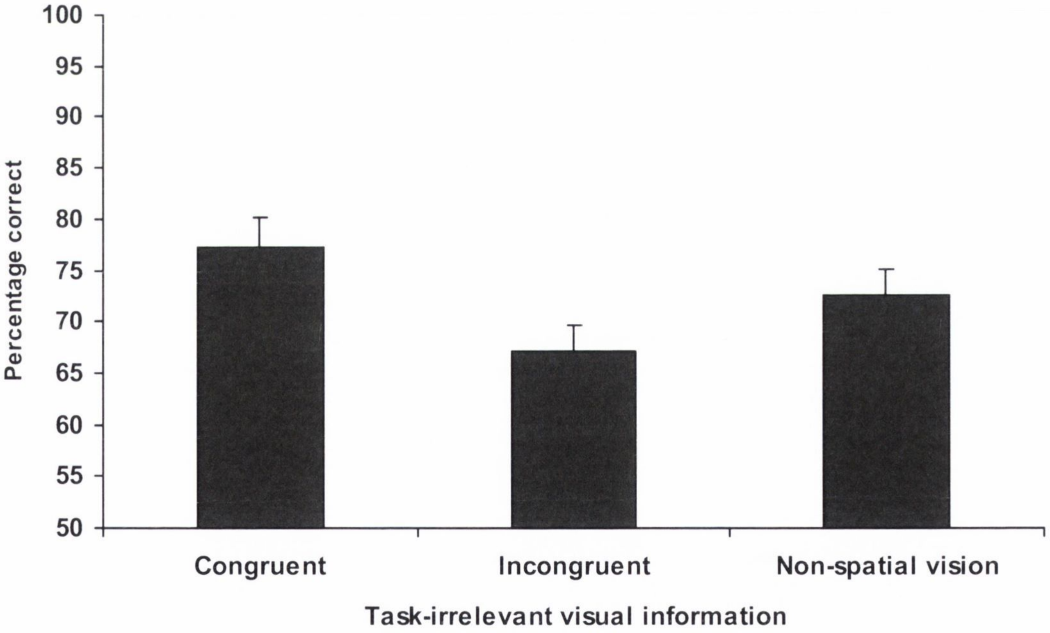

Figure 23: Plot illustrating the mean percentage correct performance in across the three conditions (Experiment 9).

It was a concern that by blocking the trials in each condition participants may have adopted specific response strategies regarding the visual stimuli in each block. As visual information is known to be more reliable for spatial localisation (Alais &

Burr, 2004) participants may have relied solely on the visual information only in the congruent condition and then compared this visual sequence with the subsequent auditory sequence in the trial. This response strategy, if indeed it was adopted, may then have contributed to the facilitation in performance. In order to test whether such a response strategy was adopted the responses within each of the conditions were analysed on a trial by trial basis. The assumption here was that a shift in performance should be observed as soon as the participant realised that the visual information was congruent to the auditory sequence and then could substitute for the auditory stimuli. It was assumed that if this did occur, it was most likely to occur after the first 5 or 10 trials. As such, any shift in performance should be observable after these first few trials and, moreover, would be unique to performance in the congruent condition. An illustration of the responses in chronological order averaged across bins of 5 trials is provided in Figure 24. As can be seen, performance in the congruent condition had a similar pattern across the block and no sudden shift in performance was observed. Moreover, the pattern of performance was similar across all conditions and not particular to the congruent condition only. As such, it can assured that a response strategy shift did not occur in the congruent blocks relative to the incongruent or non-spatial visual blocks, and that the performance facilitation found in the congruent block was due to the experimental manipulations and not an artefact.

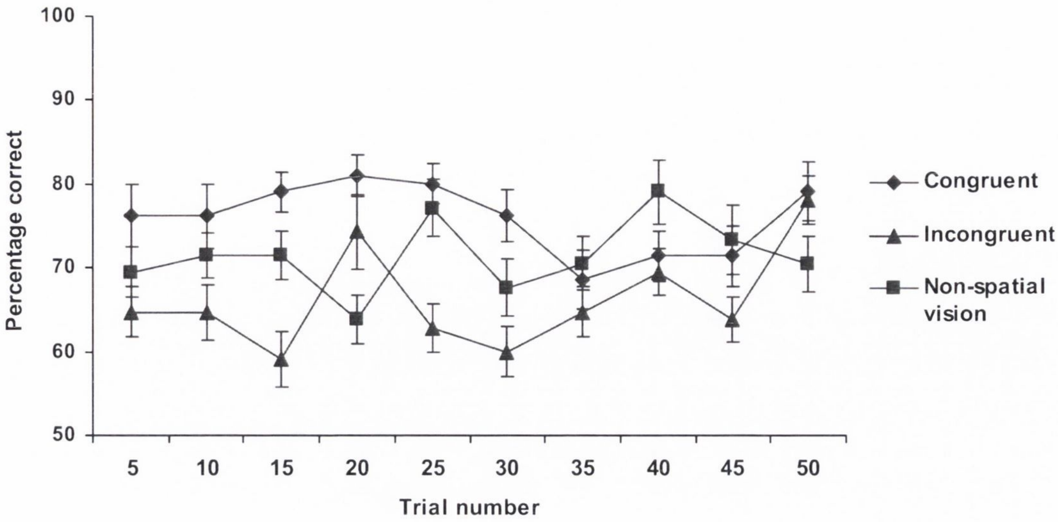

Figure 24: Plot illustrating the mean percentage correct performance in Experiment 9 plotted in chronological order across bins of 5 trials.
The mean performance is given from the start of the each block until the end (1 to 50 trials) and for each of the congruent, incongruent and non-spatial visual conditions.

5.4 General discussion

The results of Experiments 8 and 9 provide evidence for cross-modal encoding of spatial information across the senses or, more specifically, mandatory visual encoding of auditory spatial information. This study corroborates an earlier study that found a converse effect in the temporal domain (Guttman *et al.*, 2005) with auditory encoding of visual information.

The first experiment reported here demonstrated that at the temporal and spatial disparities between the auditory and visual stimuli were sufficient to avoid an effect of visual 'capture' of the auditory stimuli. Instead, the auditory and visual stimuli were perceived as being distinct from each other rendering it unlikely that visual capture explains the findings in Experiments 8 and 9. In Experiment 8 it was found that task-irrelevant visual information which was incongruent with the auditory stimuli disrupted auditory spatial processing. In Experiment 9, using a fully blocked design, a facilitation in performance was found when task-irrelevant visual information was congruent with the auditory stimuli.

In both Experiments 8 and 9, the interference stimuli were only ever present in the first auditory sequence. The findings therefore suggest that visual interference occurred at the encoding of the first auditory sequence. Moreover, these results suggest that the visual modality can influence encoded spatial information from another modality (in this case the auditory modality) when it is less reliable than the visual information.

This finding is distinct from other audiovisual spatial interactions such as the ventriloquist effect (Howard & Templeton, 1966; Jack & Thurlow, 1973), where the simultaneous presentation of visual information *alters* the perceived location of the auditory event. Under the temporal and spatial constraints used in Experiment 7 the findings indicate that the ventriloquism effect did not occur and that the auditory and visual stimuli were perceptually distinct. This finding concurs with previous research suggesting that the ventriloquism effect only occurs under specific temporal and spatial windows (Lewald *et al.*, 2001; Lewald & Guski, 2003; Radeau & Bertelson, 1987; Slutsky & Recanzone, 2001; Thomas, 1941).

The findings reported in Experiments 8 and 9 further suggest that, rather than altering the perceived location of the auditory events (i.e. visual capture), the concurrent visual information altered the encoding of this information into a visuospatial format. This result ties in with research on persons who are congenitally blind where it has been shown that the absence of visual information from birth results in both quantitative and qualitative differences in how spatial information from the intact senses is processed (Eimer, 2004; Pasqualotto & Newell, 2007; Röder, Rösler & Spence, 2004). These reports suggest a critical role for vision in setting up normal multisensory interactions for the purpose of spatial perception. This conclusion is corroborated by parallel investigations from the animal literature: the development of an accurate auditory spatial map has been shown to be dependent on normal visual experience early in life in owls (Knudsen, 1998, 2002) and in ferrets (King, 1999; King, Schnupp & Thompson, 1998).

Why vision is so important for normal spatial perception and for directing the development of neural circuits involved in spatial perception (Knudsen, 2002) has been suggested to be due to an inherent advantage of vision which is built into the genome and consequently the architecture of the brain (Witten & Knudsen, 2005). However, recent work in multisensory perception has shown that the dominance of vision is dependent on the relative reliability of the information encoded for the purpose of spatial judgements relative to any other sense (Alais & Burr, 2004; Ernst & Banks, 2002; Witten & Knudsen, 2005). This dominance in information processing could, in turn, influence the plastic changes that are required to influence the optimal development of spatial maps in other modalities (Eimer, 2004; Witten & Knudsen, 2005).

It is also possible that the influence of vision may not be to instruct the development of spatial maps at a low, unisensory level (Eimer, 2004) but it may instead interact with multisensory information at a supramodal level. In other words, modality-specific spatial maps may not exist but instead a spatial map containing spatial information from all the senses may exist from early on in spatial information processing (e.g. Wallace, Meredith, & Stein, 1992). Moreover, it may not be necessary for a supramodal cortical area to subserve a single multisensory spatial map. Instead such a map could be distributed throughout the brain and could also involve early sensory areas. Indeed, recent work has shown that multisensory interactions occur throughout what were thought to be exclusively unisensory areas (Ghazanfar & Schroeder, 2006; Schroeder & Foxe, 2005).

The results from Experiments 8 & 9 support the idea that what may be important is the spatial content of the information encoded rather than the nature of the encoding modality (i.e. vision in this case) per se. Performance in the non-spatial visual condition, which contained no relevant spatial information (as the lights were on at all spatial locations), did not differ from performance when no visual information was present (i.e. the auditory alone condition in Experiment 8), suggesting that the presence of general visual information was not sufficient to affect performance in the auditory domain. Furthermore, in comparison to performance in both the congruent and incongruent conditions, performance in the non-spatial visual condition lay mid-way between these conditions (see Experiment 9), suggesting that the effects of this visual information were relatively neutral. Both these results indicate that there is no effect on the encoding of auditory spatial perception when there was no informative spatial component in the visual information. This finding suggests that what is important is the information content rather than the sensory modality per se, since vision was not the contributing factor performance but rather the spatial information it contained.

If it is the information content rather than the sensory modality that is critical then it may be inaccurate to say that vision is involved in auditory spatial perception or that mandatory cross-modal recoding of auditory spatial information occurs. Instead, it is proposed that because auditory spatial perception involves a similar information profile as visuospatial information, it is because of these similarities in *information content* that interference then occurs. Interestingly, the idea of an amodal spatial system ties in with recent work in flavour perception: Auvray and Spence (in press) recently argued that the multisensory interactions between taste, smell, trigeminal and tactile senses (also visual and auditory cues) are combined or unified into a distinct perceptual system involved in flavour perception. Importantly, it is argued that this combination of senses involved in flavour perception is not the result of the combination of individual sensations which are then fused or bound together, but is instead the result of the multisensory information being combined or unified directly by the eating action itself. Likewise, the concerted act of orienting towards and detecting events in space could similarly unify this information from multiple modalities into a single (spatial) perceptual system (Gibson, 1966). However, further work would have to be carried out to validate this interpretation of the results.

In conclusion, task-irrelevant visual information can effect the encoding of auditory spatial information, indicating that the encoding of auditory spatial information involves vision or indeed involves spatial information which is shared between the senses.

Chapter 6. Discussion and Conclusions

In this chapter the results of my experimental studies will be discussed, and considered in respect to their implication for the existing models of synaesthesia and cross-talk theories in general. Problems these results pose for existing models will be examined and an alternative account will be sketched. Finally future research provoked by this work will be proposed.

6.1 Summary of the Results

The first three experimental chapters (Experiments 1-6) examined various aspects of linguistic-colour synaesthesia, while the last experimental chapter examined normal cross-talk in non-synaesthetes (Experiments 7-9).

The opening set of experiments (Experiment 1-3) found that synaesthetes integrated information from different senses similarly to non-synaesthetes when simple stimuli (i.e. beeps and flashes) were used but integrated in an enhanced fashion for more complex stimuli (i.e. audiovisual speech). This enhancement in audiovisual integration was only present when there was substantial noise in the visual signal, indicating that synaesthetes under noisy conditions benefit more from lip reading. Experiment 4 found that in several linguistic-colour synaesthetes that the synaesthetic colour matched the audiovisual speech percept rather than the actual audio presented (e.g. when the audio word ‘bait’ is dubbed onto a visual (viseme) of ‘gate’, audiovisual integration results in the perception of the word ‘date’, colour was induced to ‘date’ and not either of the sensory inputs). This was a consistent finding in that the synaesthetic colour never matched the actual audio when the McGurk illusion was experienced, indicating that early sensory (acoustic) information is not tied to the synaesthetic colour, but that later perceptual processes (post-integration) are critical. The final set of experiments on synaesthetes (Experiments 5-6) explored this phenomenon further, providing evidence that auditory (speech sounds) inducers are not recoded into a visual, i.e. graphemic, form before the synaesthetic colour is induced. The synaesthetic colour (as measured by the synaesthetic congruency effect) was unaffected by disruptions in grapheme processing (or general visual information), after presentation of auditory inducers indicating that the synaesthetic colour induction appears to be independent of intervening modulations of graphemic processing or basic visual processing (i.e. white noise images).

The final set of experiments (Experiments 7-9) showed that even with a task involving unimodal stimuli that there exist cross-modal effects which depend on the nature of the sensory information, which might be known as a case of pseudo-synaesthesia (Guttman *et al.*, 2005). More specifically it was shown that vision (or visuospatial information) influences the encoding of an auditory spatial stimulus even when the visual information was ignored and not relevant to the task. The reason why vision influences auditory spatial perception is because vision is more reliable at spatial information than audition. This result highlights that what is critical to information transfer or dialogue between the senses is the nature of the information involved.

6.2 Discussion of Overall Results

The first question that arises from the set of experiments examining multisensory integration in synaesthetes (Experiments 1-3) is why speech stimuli showed enhanced integration (albeit when the signal to noise ratio was low) whereas basic stimuli such as beeps and flashes showed normal integration (both in the degree of auditory capture and the amount of integration of different temporal disparities). As discussed in Chapter 2 there is an abundance of evidence in support of the idea that integration for complex speech stimuli and basic stimuli involve completely different cortical areas, so it appears that in synaesthetes there is increased cross-talk (likely reflecting structural or functional connectivity differences) between the cortical areas involved in audiovisual speech perception but not between low-level cortical areas involved in basic beeps and flashes integration. This is evidence for a broader or endo-phenotype in synaesthetes. A number of explanations (or types of explanations, see Chapter 2) may explain this differential effect for speech stimuli over more lower-level stimuli, but I would argue that it is the fact that speech stimuli are linguistic entities that may provide a clue to this difference (see part 6.4).

Another finding from these experiments is that there was no evidence for a widespread enhanced integration or hyper-integration in synaesthetes. As a number of researchers have suggested, synaesthesia may be due to a ‘hyperbinding’ mechanism most likely located in the intraparietal sulcus (Esterman *et al.*, 2006; Hubbard, 2007; Robertson, 2003; Weiss *et al.*, 2005). It would seem reasonable that if a postulated ‘binding’ mechanism was over active that this would affect all manner of ‘binding’ (or integration as these two words are often used to mean the same thing), but this was not found in the results, as synaesthetes had normal ‘binding’ of basic audiovisual stimuli.

The next experiments (Experiments 4-6) carried out on synaesthetes showed that the synaesthetic colour in linguistic-colour synaesthesia appears to be independent of early (acoustic) sensory information (Experiment 4) and also early visual information (including graphemes) (Experiment 5-6). This suggests, as with numerous other studies (Dixon *et al.*, 2000; Myles *et al.*, 2003; Rich & Mattingley, 2003; Simner & Ward, 2006), that the synaesthetic concurrent is not tightly coupled with low-level sensory properties of the inducing stimulus but that instead modulations at the perceptual level *are* closely tied to the synaesthetic colour. To reconcile some of these higher-level effects on the synaesthetic colour with accounts where the lower sensory cortices are critical (i.e. a visual grapheme cortical area), it is suggested that feedback from higher-level cortical areas modulate the pattern of activation in such lower sensory cortices which then in turn cross-activate the colour area (Hubbard & Ramachandran, 2005). This is a plausible mechanism, however the findings from Experiment 5 & 6 showed that modulations of grapheme processing did not affect the synaesthetic colour as such an account would predict. A more likely interpretation is that higher-level properties of the inducing stimulus are coupled directly to the synaesthetic colour and that no indirect route through grapheme processing areas is required. It may be important to point out that this may not have always been the case: the synaesthetic associations may have started off tightly coupled to the visual form of the grapheme but during the course of early childhood this tight coupling may have developed to a more abstract coupling. This could explain why graphemes play such a critical role in determining the synaesthetic colour of a word (Simner, Glover *et al.*, 2006) but why low-level sensory features of such are not necessary to initiate a synaesthetic colour (Dixon *et al.*, 2000; Myles *et al.*, 2003; Simner & Ward, 2006).

The experiments (Experiments 7-9) examining normal cross-talk between the senses showed that encoding (or extraction) of auditory spatial information either involved vision or, as it was argued, involved an amodal spatial perceptual system. The amodal approach to multisensory interactions follows on from J. J. Gibson's work (Gibson, 1966), where he argued that the senses should not be considered as separate and passive producers of sensations, with the separation being based on their sensory transduction properties. Instead he argued that the senses should be considered as active perceptual systems, involved in the extraction of information in order to enable an organism to function appropriately in its environment. This information extraction cuts across sensory channels or receptors types, and thus interactions between the senses are not thought as the interaction or 'binding' of sensations (or indeed representations) produced by the sensory channels but instead thought of as direct extraction of invariant information across multiple sources.

Although this is not the standard way to think about multisensory perception or how the senses communicate, recently such an approach has been applied to flavour perception (Auvray & Spence, in press) i.e. the interactions between taste and smell are combined or unified into one perceptual system concerned with flavour and thus the process of ingestion. It has been argued (Stevenson, & Boakes, 2004; Stevenson, & Tomiczek, 2007) that synaesthesia is similar to flavour perception in that it may be a synaesthesia everyone exhibits, i.e. certain odors presented alone can induce a 'synaesthetic' experience of taste which is common to all. However, this example is based on a modal approach to flavour perception and synaesthesia, in that sensations, sensory impressions or patterns of sensory stimulation from one modality cross-activate sensory impressions in another. It has been argued that it is more informative to apply an amodal approach to flavour perception (Auvray & Spence, in press), raising the question of whether this is also the case for synaesthesia. In an amodal approach any unification or cross-talk between the senses is the result of information from each sensory channel being used by the organism for a specific action in the world. With flavour perception this action is the control of ingestion and with the suggested spatial perceptual system (Chapter 5) what unifies information from multiple modalities is the action of orientating towards and identifying events in space. Pseudo-synaesthetic interactions observed in both flavour and space perception (or time perception (Guttman *et al.*, 2005)) is due on this account, to the interactions (at an informational level) within the single perceptual systems. Synaesthesia though is often thought of as a neural connectivity variant or quirk, with a completely arbitrary cross-association, and thus not due to a deviation in some higher-level system (perceptual or otherwise). However, it has been shown though that synaesthetes and non-synaesthetes share many similarities in the associations they make (Sagiv & Ward, 2006; Smilek *et al.*, 2007; Ward *et al.*, 2006), suggesting that synaesthetes may exhibit an exaggeration of some process or system that is common to us all. Previously it has been suggested that the exaggerated process is the integration process itself (Esterman *et al.*, 2006; Robertson, 2003; Ward *et al.*, 2006) or a process linking perception and linguistic thought (Simner & Ward, 2006). However, in the following sections I will consider another possibility.

6.3 Implications for Neural Models

All contemporary models agree that synaesthesia is the result of cross-talk (either through structural or functional differences) between one cortical area to another (Hubbard & Ramachandran, 2005, Ward & Mattingley, 2006). This cross-talk can be feedback from a higher level cortical areas either involved in meaning (Dixon *et al.*, 2000) or in integrating information (Esterman *et al.*, 2006; Muggleton *et al.*, 2007; Weiss *et al.*, 2005) to the area involved in processing the concurrent (e.g. colour area hV4), or it can be direct cross-activation of two adjacent cortical areas (adjacent areas in the fusiform gyrus involved in grapheme and colour processing) (Ramachandran & Hubbard, 2001b). The results of Experiments 4-6 suggest that low level sensory information or graphemic information are not tied to the synaesthetic concurrent for linguistic-colour synaesthesia, instead the synaesthetic association appears to be manifested at the perceptual level or a more higher-level of information processing. These results do not support the idea of direct cross-activation between a grapheme area and a colour area, or as discussed already, they do not support modifications of this theory which suggest that higher level interactions occur due to feedback to lower-level areas involved with the inducer.

The results of the studies carried out here do appear to support what could be called 'late' acting theories of synaesthesia i.e. where feedback from some higher level function (conceptual information or 'binding') is implicated. However, I have argued that an all purpose hyperbinding mechanism is implausible giving the results of Experiment 1-3 and what we know about multisensory integration, particularly if you take an amodal approach[1] (Gibson, 1966). Other 'late' acting theories emphasis the role of meaning (Dixon *et al.*, 2000) or linguistic thought (Simner & Ward, 2006) in linguistic-colour synaesthesia. The studies carried out in this thesis do not provide evidence against such theories (or indeed for these theories). However, it has been argued that through simulation of sensory or perceptual information many aspects of cognition occur (Barsalou, 1999; Clark, 1997; Hesslow, 2002; Hurley, 2005) so the idea of a 'concept area' stimulating a colour area would not sit well with these positions, given that they would argue that such a concept area does not exist.

It may indeed be worth considering an alternative explanation for synaesthesia, one that does not postulate cross-talk from one cortical area to another or from one sensory map to another. The situation in synaesthesia may be the result of

---

[1] Also from a philosophical perspective the idea of 'binding' has been severely criticised (Dennett, 1991; Dennett & Kinsbourne, 1992; Noë, 2005; O'Regan and Noë, 2001) as it is claimed it makes the mistake of thinking that because the world appears bound or unified in our experience that this too is the way it has to be in brain i.e. in an internal representation. This view has to deal with who 'views' this internal representation (a homunculus?) or where it is 'presented' (Dennett's Cartesian theatre?) (Dennett, 1991) and also why the creation of such a unified internal representation would explain why perceptual experience is unified (O'Regan & Noë, 2001).

a deviation in some other neurodevelopmental mechanism or could be due to malfunctioning of some overarching functional organisation, and thus not be down ultimately to aberrant connectivity (structural or functional) between two maps (sensory or conceptual) in the brain. This is not to say that there is no connectivity differences in synaesthetes, as it has been shown in a recent study that there is structural differences in synaesthetes in various parts of their brains (Rouw & Scholte, 2007), it is just making the point that this extra connectivity may not be playing the role of map to map cross-activation, and even if it is, this may not be ultimately the cause of synaesthesia but could indeed be the manifestation of some other anomaly.

To give an example, it is possible that a synaesthetic-like phenomenon could result from an imbalance in plasticity. This unexplored possibility has been implicated in many disorders (Peled, 2005) and particularly in explanations of schizophrenia (Guterman, 2006; Stephan, Baldeweg & Friston, 2006), where it is argued that hypoplasticity can result in a stabilisation of patterns of connectivity which are strongly resistant to change and hyperplasticity results in very impressionable neural networks resulting in unstable patterns of activation (Guterman, 2006). Synaesthesia could arise, according to this framework, as a regional specific hypoplasticity between two cortical areas, with the result that certain learned or encountered associations are maintained throughout adulthood. In such a scenario hyperconnectivity in synaesthetes may be observed but the source of this hyperconnectivity would not be a defective axonal guidance mechanism but instead a defective plasticity mechanism. However, any theory that postulates a plasticity imbalance (or indeed a functional imbalance) has to explain how specificity is generated (i.e. why certain areas are affected and others are not), whereas an axonal guidance mechanism has specificity built-in given its function (Dodd & Jessell, 1988).

6.4 Is language Special in Synaesthesia?

The results of experiment 1-3 demonstrated that synaesthetes showed normal integration of basic stimuli but for speech stimuli synaesthetes showed enhanced integration under noisy conditions. The suggestion was made that this may be because speech stimuli are linguistic and that synaesthesia and linguistic processes may be more intertwined than the common idea of synaesthesia as a ‘merging of the senses’ implies (Chapter 2, see also Simner, 2007). Indeed, the types of synaesthesia which contain linguistic inducers have been approximated to be as high as 88% (Simner, Mulvenna *et al.*, 2006, Simner, 2007) of all types of synaesthesia. Linguistic inducers can elicit many different types of synaesthetic concurrents; colour, spatial arrangements (Smilek *et al.*, 2006), tastes (Ward & Simner, 2003), smells (Ward *et al.*, 2005) or even personalities and gender associations. Evidence exists that almost all aspects of linguistic entities can be inducers; graphemes, phonemes, morphemes, words, lexical stress and lexical semantics (see Simner, 2007 for a comprehensive review). Taken this into the consideration, the common held belief that synaesthesia is an indiscriminate sensory map to map cross-activation may be very misleading. Given that both scientists and lay peoples first response when they hear about synaesthesia is “is it not just learned associations during childhood?” it is understandable that the focus of synaesthesia research has been more concerned with its sensory or perceptual qualities rather than why linguistic elements are so predominately involved. Also, the evidence that indicates that all types of synaesthesia may be related by a singular developmental mechanism (Barnett *et al.*, 2007) begs the question of why linguistic elements are so often involved all the more significant.

This linguistic trend may be an actual outcome of the acquisition of linguistic skills themselves, where the developmental and cognitive processes involved (somehow) increase the likelihood of synaesthetic associations being maintained or created in the first place (as mentioned in Chapter 2 the linguistic trend may also be simply due to genetic or developmental reasons which only arbitrarily involve language). This suggestion gives a primary role for linguistic development in the setting up of synaesthetic associations. Dealing with labels or graphemes involves complex multisensory information where perceptual links or associations are explicitly made. For example, learning to read involves linking phonetic sensory information to visual arbitrary forms (i.e. graphemes) with the result that proficient readers create in themselves a sort of pseudo-synaesthesia; visual graphemic forms elicit sounds in the mind so to speak, which with practice becomes involuntary. This multisensory system (indeed it may be fruitful to consider reading as carried out by a perceptual system itself) requires a complex functional organisation which may work in parallel with an underlining hyperconnectivity and facilitate the formation of synaesthetic associations. An example of how a specific higher-level functional organisation or process may facilitate synaesthetic associations is discussed in the next section.

6.5 Perceptual Anchors

A recent study on a large subpopulation of dyslexics, those with learning disabilities, showed that dyslexics benefit less from repeated auditory stimuli than non-dyslexics (Ahissar *et al.*, 2006). Specifically, in an auditory comparison task (e.g. which tone was the highest pitch?), dyslexics are impaired (relative to controls) when the task employs a small stimuli set rather than a large one. This finding was found for both auditory tones and speech stimuli. The authors suggest that this subpopulation of dyslexics fail to form a memory trace or perceptual anchor of a repeated stimulus. On the other hand, non-dyslexics can form such perceptual anchors and that this improves their performance when the set of stimuli are small i.e. when perceptual anchors would be useful for the task. Furthermore, the authors claim that previous findings that were argued to suggest low level deficits in phonological processing in fact suggest a manifestation of this higher-level or cognitive deficit. To give one example, the most common phonological deficit is in frequency discrimination, the authors claim that the deficits found in such processing are related to the standard task employed to measure frequency discrimination i.e. comparisons with a repeated tone presented in every trial (Ahissar *et al.*, 2006). This task involves forming a perceptual trace and retaining it in memory for comparison with the next stimulus, precisely the higher level process they have found to be deficient in the dyslexic population studied. Another interesting point which the authors make is that evidence exists that this failure to form a perceptual anchor in dyslexia is not unique to the auditory modality but also occurs in the visual modality, suggesting that this process operates at an amodal level.

This divergence into dyslexia research serves a double role. First it provides an example of process which may be hyperactive in synaesthetes, and second it shows that what were originally thought to be low-level sensory anomalies may in fact be a manifestation of higher-level function or a functional organisation that has gone wrong. This is not to say that this higher-level cognitive function does not itself depend on the correct operation of lower-level sensory functions. It may very well depend on such functions, but just that without such a higher-level functional explanation, making sense of lower-level functional anomalies may result in misguided or incorrect interpretations.

Is synaesthesia a case of an exuberant perceptual anchoring system? The high prevalence of linguistic inducers and the number of diverse concurrents that linguistic inducers can induce, suggest that there is something special about language and its ability to form synaesthetic associations. One such ability which has been shown to be critical for normal language and reading skills is this ability to form perceptual anchors: if this process was hyperactive in synaesthetes it may result in extra perceptual attributes being attached to linguistic entities. For example, labelling graphemes requires the operation of a perceptual anchoring system in that two phonological memory traces have to be consolidated with a visual form i.e. the name of the symbol (p = pee) and the sound it refers to (p = puh), (curiously, it is only linguistic entities, as far as I am aware, that possess this dual perceptual association). The study reported above (Ahissar *et al.*, 2006) however was not measuring or looking at the consolidation of such perceptual anchors but it is conceivable that this process may be the same for synaesthetes and non-synaesthetes (as everyone has consolidated the perceptual anchors required for reading) but where the difference lies may be in the generation in the first of place of the perceptual anchors.

From a development perspective such a scenario is clearly a possibility given that dyslexia itself is thought to be the result of abnormal connectivity, in that all the candidate genes identified in dyslexia thus far are involved in axonal guidance and neural migration (for recent reviews see Galaburda, LoTurco, Ramus, Fitch & Rosen, 2006; McGrath *et al.*, 2006; Shastry 2007), Since these studies suggest that aberrant connectivity can disrupt such a higher-level cognitive process, presumably disruptions in the opposite direction should also be possible.

6.6 Future Directions

The studies reported here showed that in linguistic-colour synaesthesia lower-level sensory information was not tied to the synaesthetic colour, whether this was low-level acoustic information (Experiment 3) or low level visual (including grapheme) information (Experiment 4). Both these results highlight that synaesthesia is tied to higher-level properties of the inducing stimuli but what exactly these higher-level properties are would be an avenue worth investigating. Indeed, it was argued that it was the linguistic properties themselves that may facilitate synaesthetic associations, given, among other factors the high prevalence of linguistic synaesthesia. An interesting research question would be whether other less investigated and apparently non-linguistic types of synaesthesia, such as for example music-colour synaesthesia, depend on linguistic entities for their induction (or origination). Recent evidence suggests this may be the case (Ward, Tsakanikos & Bray, 2006; but c.f. de Thornley Head, 2006).

Experiment 1-3 found increased cross-talk between audiovisual areas involved in speech perception under noisy conditions. It would be interesting to carry out a similar experiment but with many different signal-to-noise ratios for both the auditory and visual stimuli. Such an experiment would enable a more accurate understanding of where synaesthetes deviate from non-synaesthetes in their degree of audiovisual speech integration. In Experiments 7-9 it was shown that vision interferes with auditory spatial perception even if it is to be completely ignored, an amodal interpretation of these results was favoured but such an account was not tested directly. An experiment worth doing would be one that pits an amodal account against a modal account, using a similar auditory localisation task.

In the last few sections of this discussion it was argued that an explanation of synaesthesia may indeed involve a higher-level component (working on-top of altered connectivity), which in the end (i.e. when there is a well supported theory of synaesthesia) would do most of the explanatory work. A possible higher-level function which could be hyperactive in synaesthetes was discussed i.e. the formation of perceptual anchors. It would be interesting to test this directly using a similar paradigm used in the Ahissar and others (2006) study. Also, electrophysiological studies could similarly address this issue using the mismatch negativity response (Ahissar *et al.*, 2006). A mismatch negativity response is where there is an electrophysiological response by the auditory cortex to an oddball stimulus placed among identical stimuli and detection of the oddball stimulus involves the creation of a perceptual trace (or perceptual anchor) and thus it provides a measure of this capacity to form perceptual anchors.

6.7 Synopsis

In conclusion, the series of studies reported in this thesis provided evidence that synaesthesia is not tied to low-level properties of the inducer but instead higher-level properties of the inducer appear critical. It was argued that this higher-level effect may be specific to linguistic processes and interestingly it was found that synaesthetes did show more integration for linguistic stimuli (speech). For cross-talk theories (i.e. multisensory integration) in general it was argued, from the results of Experiment 7-9, that an amodal framework (such as that developed for flavour perception (see Auvray & Spence, in press)) may indeed be the most informative way to deal and think about spatial and other multisensory interactions. This amodal framework influenced the dismissal of various theories of synaesthesia e.g. the hyperbinding hypothesis, and also brought into question the idea of cross-talk between adjacent areas as the only explanation of synaesthesia worth considering.

An alternative account of how synaesthetic associations could develop with out any appeal to map-to-map cross-activation (at least without the necessity of such) was suggested, in that synaesthesia may be tied to the higher-level capacity of forming perceptual anchors. Such an account provides a reason of why linguistic entities are so dominant in synaesthesia, a necessary requirement of any theory of synaesthesia. Future work is needed to test this theory or indeed possibly related theories of this fascinating condition that is synaesthesia.

References

Ahissar, M., Lubin, Y., Putter-Katz, H., & Banai, K. (2006). Dyslexia and the failure to form a perceptual anchor. *Nature Neuroscience, 9*, 1558-1564.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14*, 257–262.

Aleman, A., Rutten, G. J., Sitskoorn, M. M., Dautzenberg, G., & Ramsey, N. F. (2001). Activation of striate cortex in the absence of visual stimulation: an fMRI study of synesthesia. *Neuroreport, 12*, 2827-2830.

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

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research, 21*(3), 301-8.

Andersen, T. S., Tiippana, K., & Sams, M. (2005). Maximum likelihood integration of rapid flashes and beeps. *Neuroscience Letters*, *380*, 155–160.

Ansari, D., & Karmiloff-Smith, A. (2002). Atypical trajectories of number development: a neuroconstructivist perspective. *Trends in Cognitive Sciences, 6*, 511–516.

Auvray, M., & Spence, C. (in press). The multisensory perception of flavour. *Consciousness and Cognition* (2007)*,* doi:10.1016/j.concog.2007.06.005

Ayres, A. J. (1979). *Sensory Integration and the Child.* Los Angeles: Western Psychological Association.

Barnett, K. J., Finucane, C., Asher, J. E., Bargary, G., Corvin, A. P., Newell, F., & Mitchell, K. J. (2007). Familial patterns and the origins of individual differences in synaesthesia. *Cognition*, [Epub ahead of print].

Barnett, K. J., & Newell, F. N. (2007). Synaesthesia is associated with enhanced, self-rated visual imagery. *Conscious and Cognition,* [Epub ahead of print].

Barnett, K. J., Foxe, J. J., Molholm, S., Kelly, S. P., Shalgi, S., Mitchell, K. J., & Newell, F. N. (in preparation). Differences in early sensory-perceptual processing in synesthesia: a visual evoked potentials study.

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia: Prevalence and familiality. *Perception, 25*(9), 1073–1080.

Baron-Cohen, S., Wyke, M. A., & Binnie, C. (1987). Hearing words and seeing colours: An experimental investigation of a case of synaesthesia. *Perception, 16*(6), 761–767.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-609.

Beeli, G., Essien, M., & Jancke, L. (2005). Synaesthesia: when coloured sounds taste sweet. *Nature, 434*, 38.

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception and Psychophysics*, *29*, 578–584.

Bhattacharya, J., Shams, L., Shimojo, S. (2002). Sound-induced illusory flash perception: role of gamma band responses. *Neuroreport, 13*(14), 1727-30.

Blakemore, S. J., Bristow, D., Bird, G., Frith, C., & Ward, J. (2005). Somatosensory activations during the observation of touch and a case of vision-touch synaesthesia. *Brain*, *128*, 1571–1583.

Bresciani, J. P., Dammeier, F., & Ernst, M. O. (2006). Vision and touch are automatically integrated for the perception of sequences of events. *Journal of vision*, *6*, 554–564.

Burrack, A., Knoch, D., & Brugger, P. (2005). Mitempfindung in synaesthetes: Coincidence or meaningful association? *Cortex, 42*(2), 151–154.

Callan, A. M., Callan, D. E., & Masakia, S. (2005). When meaningless symbols become letters: Neural activity change in learning new phonograms. *NeuroImage, 28*, 553-562

Calvert, G. A., Brammer, M. J., Bullmore, E. T., Campbell, R, Iversen, S. D., & David, A. S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport, 10*, 2619–23.

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*, 649-57.

Campbell, C., & Massaro, D. (1997). Perception of visible speech: Influence of spatial quantization. *Perception, 26*, 129–46.

Cannon, T. D., & Keller, M. C. (2006). Endophenotypes in the genetic analyses of mental disorders. *Annual Review of Clinical Psychology, 2*, 267-90.

Canon, L. K. (1970). Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *Journal of Experimental Psychology, 84*, 141–147.

Cerf-Ducastel, B., Van de Moortele, P. F., MacLeod, P., Le Bihan, D., & Faurion, A. (2001). Interaction of gustatory and lingual somatosensory perceptions at the cortical level in the human: a functional magnetic resonance imaging study. *Chemical Senses, 26*, 371-383.

Cherniak, C. (1994). Component placement optimization in the brain. *Journal of Neuroscience, 14*, 2418-2427.

Cherniak, C., Mokhtarzada, Z., Rodriguez-Esteban, R., & Changizi, K. (2004). Global optimization of cerebral cortex layout. *Proceedings of the National Academy of Sciences of the United States of America, 101*(4), 1081-6.

Chklovskii, D. B., Schikorski, T., & Stevens, C. F. (2002). Wiring optimization in cortical circuits. *Neuron, 34*(3), 341-7.

Clark, A. (1997). Being There: Putting Mind, Body and World Together Again. Cambridge, MA: MIT Press.

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: the case for the visual word form area. *Neuroimage, 22*, 466–76.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain, 125*, 1054-1069.

Cohen Kadosh, R., & Henik, A. (2007). Can synaesthesia research inform cognitive science? *Trends Cognitive Sciences, 11*(4):177-84.

Cohen Kadosh, R., Sagiv, N., Linden, D. E., Robertson, L. C., Elinger, G., & Henik, A. (2005). When blue is larger than red: colors influence numerical cognition in synesthesia. *Journal of Cognitive Neuroscience, 17*(11), 1766-73.

Colavita, F. B., & Weisberg, D. (1979). A further investigation of visual dominance. *Perception and Psychophysics, 25*(4), 345-7.

Cytowic, R. E. (1989). Synesthesia and mapping of subjective sensory dimensions. *Neurology, 39*(6), 849-50.

Cytowic, R. E. (2002). *Synaesthesia: A union of the senses* (2nd edition ed.). Cambridge, MA: MIT Press.

Day, S. (2005). Some demographic and socio-cultural aspects of synesthesia. *In Synesthesia: Perspectives from Cognitive Neuroscience*, L. Robertson and N. Sagiv, eds. (Oxford: Oxford University Press), pp. 11–33.

De Gelder, B., Vroomen, J., Annen, L., Masthof, E., & Hodiamont, P. (2002). Audio-visual integration in schizophrenia. *Schizophrenia Research, 59*, 211-218

De Thornley Head, P. (2006). Synaesthesia: pitch-colour isomorphism in RGB-space? *Cortex, 42*(2), 164-74

Dehaence, S., Leclec'h, G., Poline, J-B, Lebihan, D., & Cohen, L. (2002). The visual word form area. A prelexical representation of visual words in the fusiform gyrus. *NeuroReport, 13*, 321-5,

Dekle, D. J., Fowler, C. A., & Funnell, M. G. (1992). Audiovisual integration of real words. *Perception and Psychophysics, 52*, 355–62.

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

Dennett, D.C. (1991). Consciousness Explained. Boston: Little, Brown, and London: Allen Lane.

Dietz, N. A., Jones, K. M., Gareau, L., Zeffiro, T. A., & Eden, G. F. (2005). Phonological decoding involves left posterior fusiform gyrus. *Human Brian Mapping, 26*, 81-93.

Dindia, K., & Allen, M. (1992). Sex differences in self-disclosure: A meta-analysis. *Psychological Bulletin, 112*(1), 106–124.

Dixon, M. J., Smilek, D., Duffy, P. L., Zanna, M. P., & Merikle, P. M. (2006). The role of meaning in grapheme-colour synaesthesia. *Cortex. 42*(2), 243-52.

Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal: Projector versus associator synaesthetes. *Cognitive, Affective & Behavioral Neuroscience*, *4*(3), 335–343.

Dixon, M. J., Smilek, D., Cudahy, C., & Merikle, P. M. (2000). Five plus two equals yellow: mental arithmetic in people with synaesthesia is not coloured by visual experience. *Nature, 406*, 365.

Dixon, M.J., & Smilek, D. (2005). The importance of individual differences in grapheme-color synesthesia. *Neuron, 45*, 821–3.

Dodd, J., & Jessell T. M. (1988). Axon guidance and the patterning of neuronal projections in vertebrates. *Science, 242*(4879), 692-9.

Eagleman, D. M., Sarma, A. K., Nelson, S., Leal, S., Milewicz, D., & Avidan, N. [abstract]. Colored sequence synesthesia: from genetics to behavior. 2007 San Diego, CA: *Society for Neuroscience.*

Edquist, J., Rich, A. N., Brinkman, C., & Mattingley, J. B. (2005). Do synaesthetic colours act as unique features for visual search? *Cortex, 42*(2), 221–231.

Eimer, E. (2004). Multisensory Integration: How Visual Experience Shapes Spatial Perception. *Current Biology*, *14*, 115–117.

Elias, L. J., Saucier, D. M., Hardie, C., & Sarty, G. E. (2003). Dissociating semantic and perceptual components of synaesthesia: behavioural and functional neuroanatomical investigations. *Brain Research Cognitive Brain Research, 16*, 232–237.

Ernst, M. O. (2005). A bayesian view on multimodal cue integration. In G. Knoblich, I. M. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Perception of the human body from the inside out* (pp. 105–131). New York, USA: Oxford University Press.

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

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Science, 8*, 162–169.

Esterman, M., Verstynen, T., Ivry, R. B & Robertson, L. C. (2006). Coming unbound: disrupting automatic integration of synesthetic color and graphemes by transcranial magnetic stimulation of the right parietal lobe. *Journal of Cognitive Neuroscience, 18*, 1570-1576.

Foucher, J. R., Luck, D., Lacambre, M., Pham, B. T., Giersch, A., & Elliott, M.A. (in press). Poorer time resolution in schizophrenia: Longer windows of simultaneity in visual, auditory and bimodal stimuli. *Schizophrenia Research*, (2007).

Fowler, C. A., & Dekle, D. J. (1991). Listening with eye and hand: Cross modal contributions to speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 816–28.

Foxe J. J., Morocz I. A., Higgins B. A., Murray M. A., Javitt D. C., & Schroeder C. E. (2000). Multisensory auditory-somatosensory interactions in early cortical processing. *Cognition and Brain Research, 10*, 77–83.

Foxe, J. J., Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport, 16*, 419-23.

Galaburda, A. M., LoTurco, J., Ramus, F., Fitch, R. H., & Rosen G. D. (2006). From genes to behavior in developmental dyslexia. *Nature Neuroscience, 9*(10), 1213-7.

Ghazanfar A. A., & Schroeder C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences, 10*, 278-85

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *The American Journal of Psychiatry, 160*(4), 636-45

Gottesman, I. I., & Shields, J., (1972). *Schizophrenia and Genetics; a Twin Study Vantage Point.* Academic Press: New York.

Green, K. P., Kuhl, P. K., Meltzoff, A. N., & Stevens, E. B. (1991). Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics, 50*, 524–6.

Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synaesthesia: physiological constraints. *Trends in Cognitive Sciences*, **5**:36-41,

Guterman, Y. (2007). A neural plasticity perspective on the schizophrenic condition. *Consciousness and Cognition, 16*(2), 400-20.

Guttman, S.E., Gilroy, L.A., & Blake, R. (2005). Hearing What the Eyes See Auditory Encoding of Visual Temporal Sequences. *Psychological Science*, *16*, 228-235.

Gray, J. (2003). How are qualia coupled to functions? *Trends in Cognitive Sciences*, 7, 192-193

Gray, J. A., Parslow, D. M., Brammer, M. J., Chopping, S., Vythelingum, G. N., & Ffytche, D. H. (2006). Evidence against functionalism from neuroimaging of the alien colour effect in synaesthesia. *Cortex, 42*(2), 309-18.

Hairston, W. D., Burdette, J. H., Flowers, D. L., Wood, F. B., Wallace, M. T., (2005). Altered temporal profile of visual-auditory multisensory interactions in dyslexia. *Experimental Brain Research, 166*, 474-80.

Hay, J. C., Pick, H. L., & Ikeda, K. (1965). Visual capture produced by prism spectacles. *Psychonomic Science, 2*, 215–216.

Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences, 6*, 242-247

Holmes, N. P., & Spence, C. (2005). Multisensory integration: space, time and superadditivity. *Current Biology, 15*(18), 762-4.

Horváth, J., Czigler, I., Winkler, I., & Teder-Sälejärvi, W. A. (2007). The temporal window of integration in elderly and young adults. *Neurobioogy of Aging, 28*(6), 964-75.

Howard, I. P., & Templeton, W. B. (1966). *Human Spatial Orientation*. London, New York: Wiley.

Hubbard, E. M. (2007). A real red letter day. *Nature Neuroscience, 10*, 671-672.

Hubbard, E. M., Arman, A. C., Ramachandran, V. S., & Boynton, G. M. (2005). Individual differences among grapheme-color synesthetes: Brain-behavior correlations. *Neuron, 45*, 975-84.

Hubbard E. M., Manohar, S., & Ramachandran V. S. (2006). Contrast affects the strength of synesthetic colors. *Cortex, 42*, 184-9.

Hubbard, E.M., Piazza, M. Pinel, P. and Dehaene, S. (2005). Interactions between numbers and space in parietal cortex. *Nature Reviews Neuroscience, 6*(6), 435-448.

Hubbard, E. M., Simner, J. & Ward, J. (2007). Anatomically constrained cross-activation: A grand unified theory of synesthesia? [abstract]. *American Synesthesia Association*, St. Petersburg, Florida.

Hubbard E. M., & Ramachandran V. S. (2005). Neurocognitive mechanisms of synesthesia. *Neuron, 48*, 509-20.

Hurley, S. (2005) 'The Shared Circuits Hypothesis: A Unified Functional Architecture for Control, Imitation and Simulation' in S. Hurley and N. Chater, eds, *Perspectives on Imitation* (Cambridge MA, MIT Press).

Hurley, S. L., & Noë, A. (2003). Neural plasticity and consciousness. *Biology and Philosophy, 18*,131–168.

Hurley, S. L., & Noë, A. (2006). "Can Hunter-Gatherers Hear Color?". In: G. Brennan, R. Goodin, F. Jackson, M. Smith (eds.), *Common Minds: Essays in honour of Philip Pettit,* Oxford: Oxford University Press.

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology, 387*, 167–178.

Huttenlocher, P. R., & de Courten, C. (1987). The development of synapses in striate cortex of man. *Human Neurobiology*, 6, 1–9.

Ishihara, S., (1994). *Ishihara's Tests for Colour Blindness*. Kanehara and Co, Tokyo.

Jack, C. E., & Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the ''ventriloquism'' effect. *Perception and Motor Skills, 37*, 967–979.

Jacobs, L., Karpik, A., Bozian, D., & Gøthgen, S. (1981). Auditory-visual synesthesia: sound-induced photisms. *Archives of Neurology, 38*(4), 211-6.

Jansari, A. S., Spiller, M. J., & Redfern, S. (2006). Number synaesthesia: when hearing "four plus five" looks like gold. *Cortex, 42*, 253-8.

Johnson, M. H. (2001). Functional brain development in humans. *Nature Reviews Neuroscience, 2*, 475–483.

Jones, J. A., & Jarick, M. (2006). Multisensory integration of speech signals: the relationship between space and time. *Experimental Brain Research, 174*(3):588-94.

Kaiser, M., & Hilgetag, C. C. (2006). Non-optimal component placement, but short processing paths, due to long-distance projections in neural systems. *PLoS Computational Biology, 2*(7), e95.

Karmiloff-Smith, A. (1998). Development itself is the clue to understanding developmental disorders. *Trends in Cognitive Sciences, 2*(10), 389–398.

Karmiloff-Smith, A. (2006). The tortuous route from genes to behavior: A neuroconstructivist approach. *Cognitive, Affective, & Behavioral Neuroscience, 6*(1), 9-17.

Keppel, G., & Wickens, T.D. (2004). *Design and analysis: A researchers handbook* (4rd Edition). Upper Saddle River, NJ: Pearson.

Krubitzer, L., & Kahn D. M. (2003). Nature versus nurture revisited: an old idea with a new twist. *Progress in Neurobiology, 70*(1), 33-52.

Kim C. Y., Blake R., & Palmeri T. J. (2006). Perceptual interaction between real and synesthetic colors. *Cortex, 42*, 195-203

Kim, C. Y., & Blake, R. (2005) Watercolour illusion induced by synaesthetic colours. *Perception, 34*, 1501-1507.

King, A. J. (1999). Sensory experience and the formation of a computational map of auditory space in the brain. *BioEssays*, *21*, 900–911.

King, A. J., Schnupp, J. W. H., & Thompson, I. D. (1998). Signals from the superficial layers of the superior colliculus enable the development of the auditory space map in the deeper layers. *Journal of Neuroscience, 18*, 9394–9408.

Klyachko, V. A., & Stevens, C. F. (2003). Connectivity optimization and the positioning of cortical areas. *Proceedings of the National Academy of Sciences of the United States of America, 100*(13), 7937-41.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neuroscience, 27*, 712-719.

Knudsen, E. I. (1998). Capacity for plasticity in the adult owl auditory system expanded by juvenile experience. *Science*, *279*, 1531-1533.

Knudsen, E. I. (2002). Instructed learning in the auditory localization pathway of the barn owl. *Nature*, *417*, 322–328.

Laeng, B., Svartdal, F., & Oelmann, H. (2004). Does color synesthesia pose a paradox for early-selection theories of attention? *Psychological Science, 15*(4), 277-81.

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

Lewald, J., Ehrenstein, W. H. & Guski, R. (2001). Spatio–temporal constraints for auditory–visual integration. *Behav. Brain Res.,* 121, 87–103.

MacDonald, J., Andersen, S., & Bachmann, T. (2000). Hearing by eye: How much spatial degradation can be tolerated? *Perception,* 29, 1155–68.

Macleod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163-203.

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: evidence for a continuum of automaticity. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 126–135.

MacSweeney, M., Amaro, E., Calvert, G. A., Campbell, R., David, A. S., McGuire, P., & Williams, S. C. (2000). Silent speechreading in the absence of scanner noise: an event-related fMRI study. *Neuroreport, 11*, 1729–33.

Marks, L. E. (1975). 'On colored–hearing synesthesia: Cross-modal translations of sensory dimensions'. *Psychological Bulletin, 82*, 303–31.

Marks, L. E., & Odgaard, E. C. (2005). Developmental constraints on theories of synesthesia. In L.C. Robertson and N. Sagiv (Eds.), *Synesthesia: Perspectives from Cognitive Neuroscience* (pp. 214-236). New York: Oxford University Press.

Mattingley, J. B., Payne, J. M., & Rich, A. N. (2006). Attentional load attenuates synaesthetic priming effects in grapheme-colour synaesthesia. *Cortex, 42*, 213-21.

Mattingley, J. B., Rich, A. N., Yelland, G., & Bradshaw, J. L. (2001). Unconscious priming eliminates automatic binding of colour and alphanumeric form in synaesthesia. *Nature, 410*, 580-2.

McGrath, L. M., Smith, S. D., & Pennington, B. F. (2006). Breakthroughs in the search for dyslexia candidate genes. *Trends in Molecular Medicine, 12*(7), 333-41.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 265*, 746–8

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology, 42*, 135-159.

Mills, C. B., Boteler, E. H., & Oliver, G. K. (1999). Digit synaesthesia: A case study using a Stroop-type test. *Cognitive Neuropsychology, 16,* 181-191,

Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience, 27*(15), 4120-31.

Mitchell, K. J. (in press). Familial patterns and the origins of individual differences in synaesthesia. *Cognition*, (2007).

Mitchison, G. A. (1992). Axonal trees and cortical architecture. *Trends in Neuroscience, 15*, 122-126.

Molholm, S., & Foxe, J. J. (2005). Look 'hear', primary auditory cortex is active during lip-reading. *Neuroreport*, 16, 123–124.

Muggleton, N., Tsakanikos, E., Walsh, V. & Ward, J. (2007). Disruption of synaesthesia following TMS of the right posterior parietal cortex. *Neuropsychologia*, 45, 1582-5

Mulvenna C. M., & Walsh, V. (2006). Synaesthesia: supernormal integration? *Trends in Cognitive Sciences, 10*, 350-2.

Mulvenna, C. M. (2007). Synaesthesia, the arts and creativity: a neurological connection. *Frontiers of Neurology and Neuroscience, 22*, 206-22.

Mulvenna, C. M., & Walsh, V. (2006). Synaesthesia: supernormal integration? *Trends in Cognitive Sciences, 10*, 350-2

Myles, K. M., Dixon, M. J., Smilek, D., & Merikle, P. M. (2003). Seeing double: the role of meaning in alphanumeric-colour synaesthesia. *Brain and Cognition, 53*, 342–5.

Noë, A. (2005). *Action in Perception*. Cambridge, MA: MIT Press.

Noë, A. & Hurley, S. L. (2003). The deferential brain in action. *Trends in Cognitive Sciences*. 7, 195-196.

Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., Morris, R. G., Bullmore, E. T., Baron-Cohen, S., and Gray, J. A. (2002). Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. *Nature Neuroscience, 5*, 371–375.

Odgaard, E. C., Flowers, J. H., & Bradman, H. L. (1999). An investigation of the cognitive and perceptual dynamics of a colour-digit synaesthete. *Perception, 28,* 651-664.

O'Regan, K., & Noë. A. (2001). A Sensorimotor Account of Vision and Visual Consciousness *Behavioral and Brain Sciences, 24*, 939-1031

Palmeri, J., Blake, R., Marois, F., & Whetsell, W. (2002). The perceptual reality of synaesthetic colours. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 4127-31.

Pammer, K. & Vidyasagar, T. R. (2005). Integration of the visual and auditory networks in dyslexia: a theoretical perspective. *Journal of Research in Reading, 28*, 320–331

Pasqualotto, A., & Newell, F.N. (2007). The role of visual experience on the representation and updating of novel haptic scenes. *Brain & Cognition, 65*, 184–194.

Paulesu, E., Harrison, J., Baron-Cohen, S., Watson, J. D., Goldstein, L., Heather, J., Frackowiak, R. S., & Frith, C. D. (1995). The physiology of coloured hearing. A PET activation study of colour-word synaesthesia. *Brain, 118*, 661-76.

Pekkola, J., Ojanen, V., Autti, T., Jaaskelainen, I. P., Möttönën, R., Tarkiainen, A., & Sams, M. (2005) Primary auditory cortex activation by visual speech: an fMRI study at 3 Tesla. *Neuroreport, 16*, 123-4

Peled, A. (2005). Plasticity imbalance in mental disorders the neuroscience of psychiatry: implications for diagnosis and research. *Medical Hypotheses, 65*(5), 947-52.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *Journal of Neuroscience, 16*, 5205-5215.

Radeau, M., & Bertelson, P. (1974). The after-effects of ventriloquism. *Quarterly Journal of Experimental Psychology, 26*, 63–71.

Radeau, M., & Bertelson, P. (1977). Adaptation to auditory–visual discordance and ventriloquism in semirealistic situations. *Perception and Psychophysics*, *22*, 137–146.

Radeau, M., & Bertelson, P. (1987). Auditory–visual interaction and the timing of inputs: Thomas (1941) revisited. *Psychological Research, 49*, 17–22.

Ramachandran, V. S., & Hirstein, W. (1998). The perception of phantom limbs: the D.O. Hebb lecture. *Brain 121*, 1603-1630.

Ramachandran, V. S., & Hubbard, E. M. (2001a) Synaesthesia: A Window into Perception, Thought and Language. *Journal of Consciousness Studies 12*, 3-36

Ramachandran, V. S., & Hubbard, E. M. (2001b). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings of the Royal Society of London, 268*, 979-83.

Recanzone, G. H. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 869–875.

Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology, 89*, 1078–1093.

Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia: implications for the role of early experience in lexical-colour associations. *Cognition, 98*, 53-84

Rich, A. N., & Mattingley, J. B. (2002). Anomalous perception in synaesthesia: a cognitive neuroscience perspective. *Nature Reviews Neuroscience, 3*, 43-52.

Rich, A. N., & Mattingley, J. B. (2003). The effects of stimulus competition and voluntary attention on colour-graphemic synaesthesia. *NeuroReport, 14*, 1793-98.

Rich, A. N., Williams, M. A., Puce, A., Syngeniotis, A., Howard, M. A., McGlone, F., & Mattingley, J. B. (2006). Neural correlates of imagined and synaesthetic colours. *Neuropsychologia, 44*(14), 2918-25.

Ro, T., Farne, A., Johnson, R. M., Wedeen, V., Chu, Z., Wang, Z. J., Hunter, J. V., & Beauchamp, M. S. (2007). Feeling sounds after a thalamic lesion. *Annals of Neurology,* [Epub ahead of print].

Robertson, L. C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience, 4*, 93–102.

Rock, I., & Victor, J. (1964). Vision and touch: an experimentally created conflict between the two senses. *Science, 143*, 594–596.

Röder, B., Rösler, F., & Spence, C. (2004). Early vision impairs tactile perception in the blind. *Current Biology, 14*, 121–124

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-colour synesthesia. *Nature Neuroscience, 10*, 792-7.

Sagiv, N., Heer, J., & Robertson, L. (2006). Does binding of synesthetic color to the evoking grapheme require attention? *Cortex, 42*, 232-42.

Sagiv, N. & Ward, J. (2006). Crossmodal interactions: lessons from synesthesia. *Progress in Brain Research, 155*, 259-71.

Sagiv, N., & Robertson, L. C. (2005). Synaesthesia and the Binding Problem. In Synesthesia: Perspectives from Cognitive Neuroscience, L. Robertson, and N. Sagiv, eds. (Oxford: Oxford University Press), pp. 90–107.

Saint-Amour, D., De Sanctis, P. , Molholm, S., Ritter, W., & Foxe, J. J. (2007). Seeing voices: High density electrical mapping and source-analysis of the multisensory mismatch negativity evoked during the McGurk illusion. *Neuropsychologia, 45*, 587-597.

Sams, M., Manninen, P., Surakka, V., Helin, P., & Kättö, R. (1998). McGurk effect in Finnish syllables, isolated words, and words in sentences: Effect of word meaning and sentence context. *Speech Communication, 26*, 75–87.

Scerif, G., & Karmiloff-Smith, A. (2005). The dawn of cognitive genetics? Crucial developmental caveats. *Trends in Cognitive Sciences, 9*(3), 126-35.

Schroeder, C. E., & Foxe, J. J. (2002). The timing and laminar profile of converginginputs to multisensory areas of the macaque neocortex. *Cognitive Brain Research, 14*, 187–198.

Schroeder, C. E., & Foxe, J. (2005). Multisensory contributions to low-level, 'unisensory' processing. *Current Opinion in Neurobiology, 15*, 454-458.

Shams, L., Kamitani, Y., & Shimojo, S., (2000). Illusions. What you see is what you hear. *Nature*, 14, 788.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*(1):147-52

Shams, L., Kamitani, Y., Thompson, S., & Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *Neuroreport, 12*(17), 3849-52.

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

Shanon, B. (2002). Ayahuasca visualizations: a structural typology. *Journal of Conscious Studies, 9*, 3-30.

Shastry, B. S. (2007). Developmental dyslexia: an update. *Journal of Human Genetics, 52*(2), 104-9.

Shipley, T. (1964). Auditory flutter-driving of visual flicker. *Science, 145*, 1328–1330.

Sidak, Z. (1967). Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *American Statistical Association, 62*, 626-633.

Simner, J. (2007). Beyond perception: synaesthesia as a psycholinguistic phenomenon. *Trends in Cognitive Science, 11*, 23-9

Simner, J., Glover, L., & Mowat, A. (2006). Linguistic determinants of word colouring in grapheme-colour synaesthesia. *Cortex, 42*, 281-9

Simner, J., & Hubbard, E. M. (2006). Variants of synesthesia interact in cognitive tasks: evidence for implicit associations and late connectivity in cross-talk theories. *Neuroscience, 143*(3), 805-14.

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., Scott, K., & Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception, 35*, 1024–1033.

Simner, J. & Ward, J. (2006). Synaesthesia: the taste of words on the tip of the tongue.

Nature, 23, 444-438.

Simner, J., Ward, J., Lanz, M., Jansari, A., Noonan, K., Glover, L., & Oakley, D. (2005). Non-random associations of graphemes to colours in synaesthetic and non-synaesthetic populations. *Cognitive Neuropsychology, 22*(8), 1069–1085.

Simon, T. J. (2007). Cognitive characteristics of children with genetic syndromes. *Child Adolescent Psychiatric Clinics of North America, 16*, 599-616.

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

Smilek D., Callejas A., Dixon M. J., & Merikle P. M. (2006). Ovals of time: Time-space associations in synaesthesia. *Conscious Cognition, 16*(2):507-19

Smilek, D., Carriere, J. S., Dixon, M. J., & Merikle, P. M. (2007). Grapheme frequency and color luminance in grapheme-color synaesthesia. *Psychological Science, 18*(9), 793-5.

Smilek, D., Dixon, M. J., Cudahy, C., & Merikle, P. M. (2001). Synaesthetic photisms influence visual perception. *Journal of Cognitive Neuroscience, 13*, 930–6.

Smilek, D., Dixon, M. J. & Merikle, P. M. (2003). Synaesthetic photisms guide attention. *Brain and Cognition, 53*(2), 364-7.

Smith, E. G., & Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry, 48*, 813-21.

Sperling, J. M., Prvulovic, D., Linden, D. E., Singer, W., & Stirn, A. (2006). Neuronal correlates of colour-graphemic synaesthesia: a fMRI study. *Cortex, 42*(2), 295-303.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. (Cambridge, Mass.: MIT Press).

Stephan, K. E., Baldeweg, T., & Friston, K. J. (2006). Synaptic plasticity and dysconnection in schizophrenia. *Biological Psychiatry*, *59*(10), 929-39.

Stevenson, R. J., & Boakes, R. A. (2004). Sweet and sour smells: Learned synaesthesia between the senses of taste and smell. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processing* (pp. 69–83). Cambridge, MA: MIT Press.

Stevenson, R. J., & Tomiczek, C. (2007). Olfactory-induced synesthesias: A review and model. *Psychological Bulletin, 133*, 294–309.

Sykes, N. H., & Lamb, J. A. (2007). Autism: the quest for the genes. *Expert Reviews in Moecular Medicinw, 9*, 1-15.

Tarkiainen, A., Cornelissen, P.L., & Salmelin, R., (2002). Dynamics of visual feature analysis and object-level processing in face versus letter-string perception. *Brain, 125*, 1125-1136.

Thomas, G. J. (1941). Experimental study of the influence of vision on sound. *Journal of Experimental Psychology, 28*, 163–177.

Thurlow, W. R., & Jack, C. E. (1973). A study of certain determinants of the 'ventriloqism effect'. *Perception and Motor Skills, 36*, 1171–1184.

Wallace, M.T., Meredith, M.A., & Stein, B.E. (1992). Integration of multiple sensory modalities in cat cortex. *Experimental Brain Research, 91*(3), 484-488.

Walters, J. T. R., & Owen, M. J. (2007). Endophenotypes in psychiatric genetics *Molecular Psychiatry, 12*, 886 – 890.

Wang, W., Datta, H., & Sussman, E. (2005). The development of the length of the temporal window of integration for rapidly presented auditory information as indexed by MMN. *Clinical Neurophysiology, 116*(7), 1695-706

Ward, J. (2007). Acquired auditory-tactile synaesthesia. *Annals of Neurology,* [Epub ahead of print].

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: to what extent does it use cross-modal mechanisms common to us all? *Cortex, 42*, 264-80.

Ward, J., Li, R., Salih, S., & Sagiv, N. (2006). Varieties of grapheme-colour synaesthesia: A new theory of phenomenological and behavioural differences. *Conscious and Cognition,* [Epub ahead of print].

Ward, J. & Mattingley, J. B. (2006). Synaesthesia: an overview of contemporary findings and controversies. *Cortex, 42*(2), 129-36.

Ward, J., & Simner, J. (2003) Lexical-gustatory synaesthesia: Linguistic and conceptual factors. *Cognition*, 89, 237-261.

Ward, J., & Simner, J. (2005). Is synaesthesia an X-linked dominant trait with lethality in males? *Perception 34*, 611–623.

Ward, J., Simner, J., & Auyeung, V. (2005). A comparison of lexical-gustatory and grapheme-colour synaesthesia. *Cognitive Neuropsychology, 22*, 28–41.

Ward, J., Thompson-Lake, D., Ely, R., & Kaminski, F. (2007). Synaesthesia, creativity and art: What is the link? *British Journal of Psychology*. [Epub ahead of print]

Ward, J., Tsakanikos, E., & Bray, A. (2006). Synaesthesia for reading and playing musical notes. *Neurocase, 12*(1), 27-34.

Watkins, S., Shams, L., Josephs, O., & Rees, G. (2007). Activity in human V1 follows multisensory perception. *Neuroimage, 37*(2), 572-8.

Watkins, S., Shams, L., Tanaka, S., Haynes, J. D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage, 31*(3), 1247-56.

Weiss, P. H., Shah, N. J., Toni, I., Zilles, K., & Fink, G. R. (2001). Associating colours with people: a case of chromatic-lexical synaesthesia. Co*rtex, 37*, 750–753.

Weiss, P. H., Zilles, K., & Fink, G. R. (2005). When visual perception causes feeling: enhanced cross-modal processing in grapheme color synesthesia. *Neuroimage, 28*, 859-68

Welch, R.B., & Warren, D.H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*, 638–667.

Welch, R.B., & Warren, D.H. (1986). Intersensory interactions. In Handbook of Perception and Human Performance (Boff, K.R., et al., eds.), pp. 25.1–25.36, Wiley.

Williams, J. H., Massaro, D. W., Peel, N. J., Bosseler, A., & Suddendorf, T. (2004). Visual-auditory integration during speech imitation in autism. *Research in Developmental Disabilities, 25*, 559–575.

Windman, S. (2004). The effect of sentence context and expectation on the McGurk illusion. *Journal of Memory and Language, 50*, 212-230

Witten, I. B. & Knudsen, E. I. (2005). Why seeing is believing: merging auditory and visual worlds. *Neuron, 48*(3), 489-96.

Witthoft, N., & Winawer, J., (2006). Synesthetic colors determined by having colored refrigerator magnets in childhood. *Cortex, 42*, 175-83.

Wollen, K. A., & Ruggiero, F. T. (1983). Colored-letter synesthesia. *Journal of Mental Imagery, 7*, 83-86.

Yaro, C., & Ward, J. (2007). Searching for Shereshevskii: what is superior about the memory of synaesthetes? *Quarterly Journal of Experimental Psychology, 60*, 681-95.