Interdisciplinary postgraduate program in **Basic and Applied Cognitive Science**

of

National and Kapodistrian University of Athens

Is A always red? Multisensory integration of synesthetic stimuli in synesthetes and non-synesthetes

Student

Miketa Arvaniti

Supervisor

Dr. Argiro Vatakis

Co-supervisor

Dr. Noam Sagiv

Dr. Argiro Vatakis

Dr. Noam Sagiv

Acknowledgments

I am deeply grateful to my supervisor, Dr. Argiro Vatakis, for her guidance, support, and advice, as well as for showing me all the stages of a research project. I want to thank Dr. Noam Sagiv for his support, trust, and excellent cooperation during my visit at Brunel University. I also wish to thank Lucille Lecoutre for her help and kindness during my stay in Brunel University. I want to thank Emmanouil Konstantinidis for his support when I need it the most and for providing me an excellent opportunity to perform part of my experiments at the Department of Cognitive, Perceptual, and Brain Sciences of the University College London. I' m very grateful to Helena Sgouramani for her constant advice and friendship.

Last but not least, I wish to thank my family and my friends for always supporting me.

This thesis is dedicated to my teacher Dr. Argiro Vatakis

for showing me this fascinating field

Table of Contents

1. Introduction 5

1.1 The binding problem 5

1.2 Spatiotemporal factors of multisensory integration 5

1.3 The congruency factor 6

1.3.a Crossmodal correspondences 6

1.4 The Assumption of Unity 7

1.5 The phenomenon of synesthesia 8

1.6 About our research 8

2. Crossmodal correspondences 10

2.1 Introduction 10

2.2 Methods 11

2.2.a Participants 11

2.2.b Apparatus & Materials 11

2.2.c Procedure 12

2.2.d Design 12

2.3 Results 13

2.4 Discussion 15

3. Synesthetic Associations 18

3.1 Introduction 18

3.2 Methods 18

3.2.a Participants 18

3.2.b Apparatus & Materials 18

3.2.c Procedure 19

3.2.d Design *19*

3.3 Results 20

3.4 Discussion 20

4. Synesthetic Associations in Naïves 21

4.1 Introduction 21

4.2 Methods 21

4.2.a Participants *21*

4.2.b Apparatus & Materials *21*

4.2.c Procedure *22*

4.2.d Design *22*

4.3 Results 23

4.4 Discussion 23

5. Conclusion 24

References 25

1. Introduction

1.1 The binding problem

Every day we depend on our senses to perceive the world around us. Most of our experiences have multisensory attributes meaning that our brain needs to bind information from different senses in order to gain a coherent and meaningful percept of those experiences. Integration of various sensory inputs increases perceptual reliability, so it is crucial for the different sense organs to cooperate and combine the right stimuli together. How does our brain “know” which information to unify and which to process separately? Research on the binding problem focuses on the understanding of the factors that lead to multisensory integration.

1.2 Spatiotemporal factors of multisensory integration

Cognitive neuroscience investigates, among other, the temporal and spatial aspects of a multisensory experience that lead to multisensory binding. Spatial and temporal coincidence of audiovisual events is known to promote multisensory integration (e.g., Jones & Jarick, 2006; Stein & Meredith, 1993) and facilitate stimulus detection when compared to unimodal stimuli (e.g., Vroomen & Gelder, 2000).

The spatiotemporal interaction of vision and audition has typically been studied using the “ventriloquist effect”. This effect is influenced by both spatial and temporal components for integration to be achieved. In regards to the spatial factor, this effect is caused by the tendency to underestimate the distance between an auditory and a visual input that occur simultaneously from different sources and results in the perception that the auditory stimulus came from, or near, the visual input. So, spatial ventriloquism results in a shift of the apparent location of the auditory stimulus toward the visual due to their temporal coincidence (Bertelson, Frissen, Vroomen, & de

Gelder, 2006). In regards to the temporal factor, people's temporal discrimination sensitivity for two visual stimuli enhances when an auditory event comes before the first visual stimulus and another auditory event comes after the second visual stimulus. This enhancement in people's temporal sensitivity results in an expansion of the perceived interval between the two visual events (Parise & Spence, 2008).

1.3 The congruency factor

Beside the spatial and temporal factor, it is suggested that congruency can also promote integration. Congruency may have a semantic dimension or can be due to crossmodal correspondences. In order to define and study congruency in a quantitative way, this factor is usually studied by presenting an ecologically meaningful visual stimulus and a matching auditory counterpart in matched and mismatched format (e.g., Vatakis & Spence, 2007).

Semantic congruency usually refers to those situations in which pairs of auditory and visual stimuli are presented varying in terms of their identity and/or meaning (Spence, 2011). Semantic congruency has been reported for faces and gender-matched or -mismatched speech sounds (Vatakis & Spence, 2007), for object pictures (e.g., the picture of a dog), and semantically-matched or -mismatched environmental sounds (e.g., the sound of a dog) (Chen & Spence, 2010), and for letters and matched or mismatched speech sounds (e.g., van Atteveldt, Formisano, Goebel, & Blomert, 2004).

Crossmodal correspondences refer to associations between putatively nonredundant stimulus attributes or dimensions that happen to be shared by many people (Spence, 2011). Such correspondences occur usually for extreme values of a stimulus dimension, meaning that a more-or-less extreme stimulus on a given dimension is matched with a more-or-less extreme value on the corresponding dimension (Spence, 2011). For example, a large shape is associated with a low-pitched tone, while a small one with a high-pitched tone. Such correspondences are widely known as crossmodal correspondences.

1.3.a Crossmodal correspondences

Crossmodal correspondences have been studied for almost 100 years. Part of these studies has focused on sound symbolism, where certain shapes are highly associated with certain pseudowords (Köhler, 1947; Sapir, 1929). Ramachandran and Hubbard (2001; 2003) performed an experiment in which they presented one round and one angular shape and asked people which one was named “buba” and which one “kiki” (see Figure 1). Their experiment suggests that 95 to 98% of the population associates an angular shape with the word “buba” and a sharp shape with the word “kiki”. “Buba” and “kiki” shapes are considered typical examples on sound symbolism, though crossmodal correspondences exist between a variety of auditory and visual stimulus dimensions (e.g., size, brightness, loudness) and have been documented for simple stimulus dimension, such as loudness and for more complex stimuli, such as images and pseudowords.

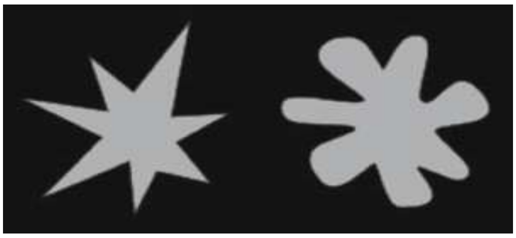

Figure 1: Ramachandran and Hubbard (2001; 2003) showed that between 95 to 98% of the population associates the angular shape on the right with the word “buba” and the sharp shape on the left with the word “kiki”.

Research on crossmodal correspondences focuses on the type of the associations, the modalities involved and their origin. Whether such correspondences are innate or learned is not clear as various studies have presented contradictory results.

Ludwig, Adachi, and Matsuzawa (2011) tried to establish whether evolutionary origins of cross-modal mappings between higher-pitched sounds with lighter colors exist. They tested whether chimpanzees (Pan troglodytes) also associate higher pitch with higher luminance. In their experimental design thirty-three humans and six chimpanzees were required to classify black and white squares according to their color while hearing irrelevant background sounds that were either high- or low-pitched. Both species performed better when the background sound was congruent (high-pitched for white, low-pitched for black) than when it was incongruent (low-pitched for white, high-pitched for black). They suggest that the tendency to pair high-pitch with high luminance evolved before the human lineage split from that of chimpanzees rather than being a culturally learned or a linguistic phenomenon.

Walker et al. (2010) examined preferential looking to assess 3- to 4-month-old preverbal infants' sensitivity to the correspondences linking auditory pitch to visuospatial height and visual sharpness. The infants looked longer at a changing visual display when this was accompanied by a sound whose changing pitch was congruent, rather than incongruent, with these correspondences. This study also indicates that crossmodal correspondences are an unlearned aspect of perception.

Towards the opposite direction, Ernst (2007) questioned the argument that the sensory system already knows which signals belong together and how they relate and suggest that a mapping between two sensory signals from vision and touch can be learned from their statistical co-occurrence such that they become integrated. He trained subjects with stimuli that are usually unrelated in the world, such as the luminance of an object (visual signal) and its stiffness (haptic signal). In the training phase, they presented subjects with combinations of these two signals, which were artificially correlated, and thus, introduced a new mapping between them. For example, the stiffer the object, the brighter it was. They measured the influence of learning by comparing discrimination performance before and after training. Their prediction was that integration makes discrimination worse for stimuli, which are incongruent with the newly learned mapping, because integration would cause this incongruency to disappear perceptually. The more certain subjects are about the new mapping, the stronger should the influence be on discrimination performance. Thus, learning in this context is about acquiring beliefs. They found a significant change in discrimination performance before and after training when comparing trials with congruent and incongruent stimuli. After training, discrimination thresholds for the incongruent stimuli are increased relative to thresholds for congruent stimuli, suggesting that subjects learned effectively to integrate the two formerly unrelated signals.

1.4 The Assumption of Unity

Both the spatiotemporal and congruency factors promote multisensory integration and consist the core of the most commonly held view on the crossmodal binding problem; the assumption of unity (e.g., Keetels & Vroomen, 2011; Vatakis & Spence, 2007). The presentation of a visual and an auditory event can be perceived either as one audiovisual event or as two separate unimodal events. Our brain "decides" whether to bind them or not based both on stimulus driven factors, such as spatiotemporal, and on cognitive factors, such as semantic congruency. The assumption of unity states that whenever there is high consistency between two or more sensory inputs, our brain is more likely to unify them and perceive them as a single multisensory event rather than as separate unimodal events (Bedford, 2001; Vatakis & Spence, 2007; Welch, 1999a, 1999b; Welch & Warren, 1980)

One way to study the "unity effect" in multisensory integration is the temporal order judgments (TOJ) task. Participants are presented with visual and auditory stimuli in various stimulus onset asynchronies (SOAs) and make judgments regarding which modality was presented first. Stimuli are presented in congruent and incongruent format. The TOJ task allow us to obtain the just noticeable difference (JND) and the point of subjective simultaneity (PSS; Vatakis & Spence, 2007). The JND provides a standardized measure of the sensitivity with which subjects judge the temporal order of the visual or auditory stimuli, while the PSS provides an estimate of the interval the visual stimulus has to lead the auditory in order for synchrony to be perceived. Congruency promotes multisensory integration, thus, it is more difficult for people to judge the order of modalities, leading to a higher JND compared to incongruent cases. This difference in the JND is known as "unity effect".

1.5 The phenomenon of synesthesia

According to the assumption of unity, as described above, people sometimes perceive two unisensory inputs as a single event, even though their perception involves different modalities (e.g., visual and auditory). In the phenomenon of synesthesia, a separate unimodal stimulus causes an atypical experience in another non-stimulated modality, leading to the perception of a single unified multisensory event. Whether the concurrent perception of two percepts coming from one sensory input in the synesthetic mind consists a case of multisensory integration remains an open question that we will attempt to approach in Chapter 3.

Synesthesia is automatic (i.e., synesthetic stimuli appear unvoluntarily; Beeli, Esslen, & Jäncke, 2008; Brang, Hubbard, Coulson, Huang, & Ramachandran, 2010; Ramachandran & Hubbard, 2001), consistent (i.e., the synesthetic experience is always the same; Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007; Simner et al., 2005), and idiosyncratic (i.e., the synesthetic experience of each synesthete is unique; Dixon & Smilek, 2005; Rouw & Scholte, 2007). There are several types of synesthesia (e.g., number-color, time-space, music-color, taste-color) involving different modality combinations, though the most studied type is grapheme-color synesthesia, where the visual and/or the auditory form of a letter induces the concurrent experience of a color.

1.6 About our research

In the present study, we aim to contribute to the current research of multisensory integration and synesthesia. We will explore the congruency factor involved in crossmodal correspondences for the dimensions of color, size, and angularity. We will also aim to determine whether there is a link between synesthesia and multisensory integration by comparing the unity effect of synesthetes and naïves, using the crossmodal correspondences mentioned above, and by measuring the unity effect of synesthetic grapheme/number-color associations. Subsequently, we will attempt to investigate whether certain grapheme-color combinations that both naïves and synesthetes make often, when they are asked explicitly, compose a case of audiovisual integration or instead are semantically driven.

2. Crossmodal correspondences

2.1 Introduction

Crossmodal correspondences occur among different senses and stimulus types. In the present experiment, we will study the integration of correspondences between vision and audition for the stimulus dimensions of color, size, and angularity. The underlying mechanism of such correspondences is still unknown, though there is growing evidence that, at least in some of the reported cases, multisensory integration is involved (Parise & Spence, 2009). Parise and Spence investigated the role of synesthetic correspondences on the integration of pairs of temporally conflicting auditory and visual stimuli. Based on the hypothesis that integration has the cost of hampering the brain's access to the individual sensory components feeding into the integrated percept, thus reducing the reliability of estimates of potential crossmodal conflicts (Ernst, 2006), they showed that for the correspondences of the stimulus dimensions of size and angularity with high- and low-pitched tones, integration is involved. They used a TOJ task with large and small, angular- and sharp-visual stimuli and low- and high-pitched tones, respectively, for each dimension. Their study showed with a psychophysical orthogonal task that congruency promotes multisensory integration.

Crossmodal correspondences are often described as a weak form of synesthesia (Martino & Marks, 2001) common to all. Also, it is suggested that synesthesia is based on universal mechanisms rather than being based on mechanisms found solely in synesthetes (Sagiv & Ward, 2006). In the present experiment, we also aimed to determine a possible link between crossmodal correspondences and synesthesia by testing synesthetes in the same pairs of crossmodal correspondences as naïves. Two recent studies have attempted to define whether synesthetes experience enhanced multisensory integration than naïves, reporting contradictory results (Brang, Williams, & Ramachandran, 2012; Neufeld, Sinke, Zedler, Emrich, & Szycik, 2012). Brang and colleagues showed that synesthetes experience enhanced multisensory integration between vision and audition. As already argued, crossmodal correspondences involve multisensory integration, thus, in order to investigate whether synesthetes are better integrators than naïves, we will test both groups in the same pairs of correspondences. We will measure multisensory integration using an unspeeded TOJ task. We expect that participants will find it harder to make a judgment regarding the modality order of the stimuli in congruent cases than in incongruent ones, leading to a higher JND value. Our hypothesis is that both synesthetes and naïves will experience a "unity effect". If synesthetes are experiencing enhanced, reduced or equal multisensory integration to naïves, we expect a higher, lower or equal JND values than naïves, respectively.

2.2 Methods

2.2.a Participants

Eighteen naïve participants (M=27 years of age) and 8 audiovisual associator synesthetes (M=28 years of age), with normal or corrected to normal vision and audition took part in this experiment. Synesthesia was confirmed using the Eagleman battery (Eagleman et al., 2007) and the associator-projector questionnaire (Skelton, Ludwig, & Mohr, 2009). All participants signed a consent form and received monetary reward for their participation.

2.2.b Apparatus & Materials

Visual stimuli were presented on a 13.3-inch computer screen (refresh rate 60Hz) against a white background. Auditory stimuli were presented via 2 speakers placed at 3 cm to either side of the computer screen. Three pairs of crossmodal correspondences were used regarding color, size, and angularity (see Figure 2). For the first pair of stimuli, the visual stimuli consisted of a square of 230x230 pixels, colored either in light yellow or dark blue. The auditory stimuli consisted of a high-and a low-pitched tone with a frequency of 4000 and 230Hz, respectively. For the second pair of stimuli, the visual stimuli consisted of a small (80 pixels) and a large circle (250 pixels) presented at the center of the screen. Auditory stimuli were the same as in the first pair. For the third pair, the visual stimuli were the shapes of “buba” and “kiki” used in Köhler’s study (1947) colored in black. Auditory stimuli consisted of the sounds of phonemes /o/ and /i/ articulated by a female voice. The congruent cases were light color-high pitched tone, dark color-low pitched tone, small circle-high pitched tone, large circle-low pitched tone, shape of “buba”-/o/, and shape of “kiki”-/i/, while the incongruent ones were consisted of the opposite pairs.

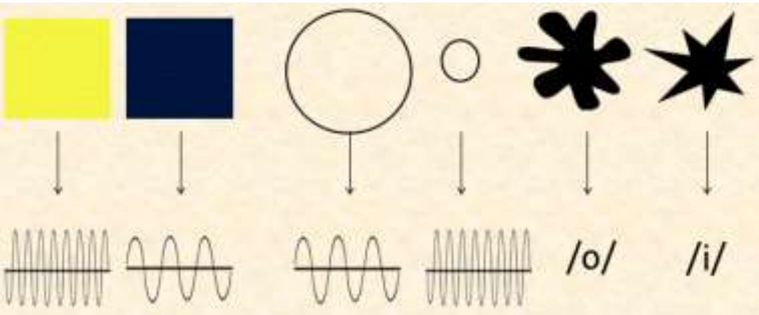

Figure 2: Samples of the stimuli used in this experiment. The stimuli were presented in congruent and incongruent format. This figure presents the congruent cases of the stimuli used in the experiment for color, size, and angularity. For example, the yellow square is paired with the high pitched sound and the blue one with the low pitched sound. Participants are expected to integrate more the yellow colored square when presented with the high pitched tone than with the low pitched. Respectively, participants are expected to experience greater integration when presented with the pairs in this figure than with their opposites.

2.2.c Procedure

Participants performed an unspeeded TOJ task regarding which stimulus was presented first (i.e., visual or auditory) by pressing one of the response keys (i.e., “V” for visual, “A” for auditory). Experimental duration was approximately 30 minutes.

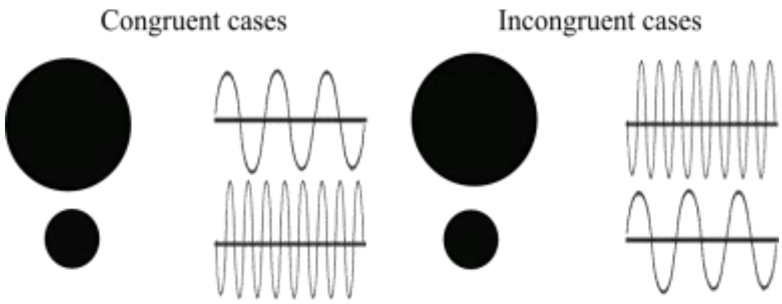

Figure 3: Each stimulus dimension (size and pitch) has 2 congruent and 2 incongruent audiovisual pairs. Each of the 4 audiovisual pairs regarding size is presented equiprobably.

2.2.d Design

The visual and auditory stimuli were presented for 200 ms in various stimulus onset asynchronies (SOAs; ±250, ±150, ±100, ±80, 0 ms; Negative values indicate that the auditory stimulus was presented first). The SOAs varied on a trial-by-trial basis using the method of constant stimuli. All pairs of audiovisual stimuli were presented randomly in congruent and incongruent format equiprobably (see Figure 3). The experiment consisted of 6 blocks of 108 trials each.

2.3 Results

Responses were collected as the proportion of “visual first” responses and converted to their equivalent z-scores, assuming a cumulative normal distribution. Data from the SOAs were used to calculate the best-fitting straight lines for each participant for each condition. These were used to derive values for the slope and the intercept. The slope and intercept values were used to calculate the JND (0.675/slope; as ±0.675 represents the 75^{th} and 25^{th} percentiles on the cumulative normal distribution) and the PSS (intercept/slope) values. For all the analyses reported here, Bonferroni-corrected t-tests (in which $P<0.05$ before correction) were used for all posthoc comparisons. Participants that their JND was greater than our largest SOA (i.e., 250ms) in one stimulus were excluded from all the analyses (n=12). One-way analysis of variance (ANOVA) with the factor of 'Stimulus' (light color congruent, dark color congruent, small circle congruent, large circle congruent, buba congruent, kiki congruent, light color incongruent, dark color incongruent, small circle incongruent, large circle incongruent, buba incongruent, kiki incongruent) didn't revealed any differences, thus the data from the pairs referring to the same dimension (i.e., color, size, and angularity) were collapsed for further analysis.

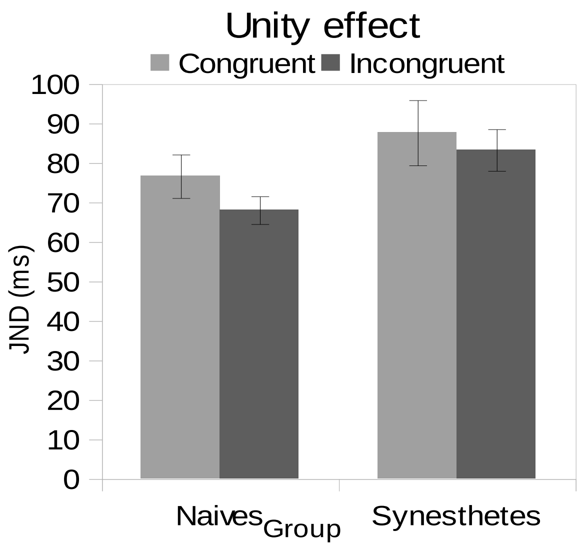

Figure 4: The JNDs in congruent cases were higher than in incongruent ones for naïves and synesthetes, suggesting a unity effect, but only for naïves this effect was significantly important. A higher JND value means that participants find more difficult to perceive the order of the modality when the stimuli "fit together".

In order to investigate whether there is a 'unity effect', we analyzed collapsed data of naïve participants using repeated-measures ANOVA with the factors of ‘Stimulus’ (color, size, and angularity) and ‘Congruency’ (congruent and incongruent). Analysis revealed a main effect of Congruency [F(1,35)=7.527, P=0.010; see Figure 4] with the JNDs in congruent cases (M=77ms) being higher than in incongruent ones (M=68ms). No main effect of Stimulus [F(2,70)=1.065, P=0.35] or interaction (Congruency*Stimulus [F(2,70)=0.125, P=0.883]) was obtained.

In terms of group differences, the JNDs of the collapsed data were analyzed using repeated-measures ANOVA with the within-subjects factors of ‘Stimulus’ (color, size, and angularity) and ‘Congruency’ (congruent and incongruent) and the between-subjects factor of 'Group' (naïve or synesthete). This analysis revealed no main effects of Stimulus, Congruency, Group or an interaction (p>0.05). The main effect of Congruency was marginal [F(1,50)=3.39, p=0.072]. Post-hoc analysis revealed significant effect of Congruency in naïves [F(1,50)=4.833, P=0.033], with the JNDs in congruent cases (M=77ms) being higher than the incongruent ones (M=68ms). In synesthetes, the JNDs in congruent cases (M=88ms) were higher than in incongruent cases

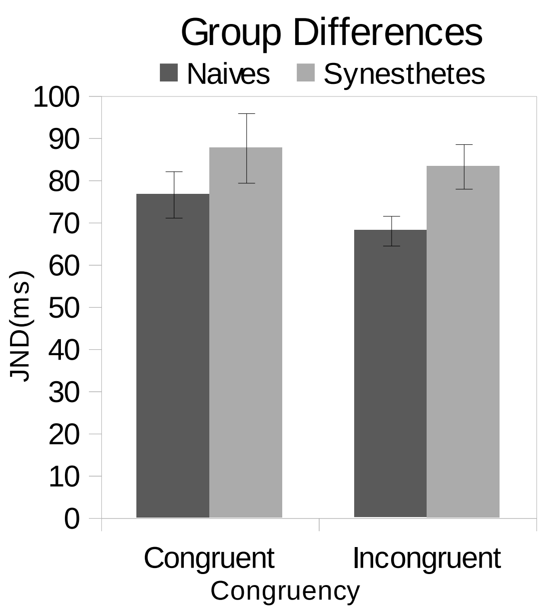

Figure 5: Synesthetes have higher JNDs both in congruent and in incongruent case than naïves, suggesting that they face greater difficulty in define the temporal order of two stimuli irrespective of congruency.

(M=83ms; see Figure 4), however this effect didn't reach significance probably due to the small number of synesthetes recruited so far.

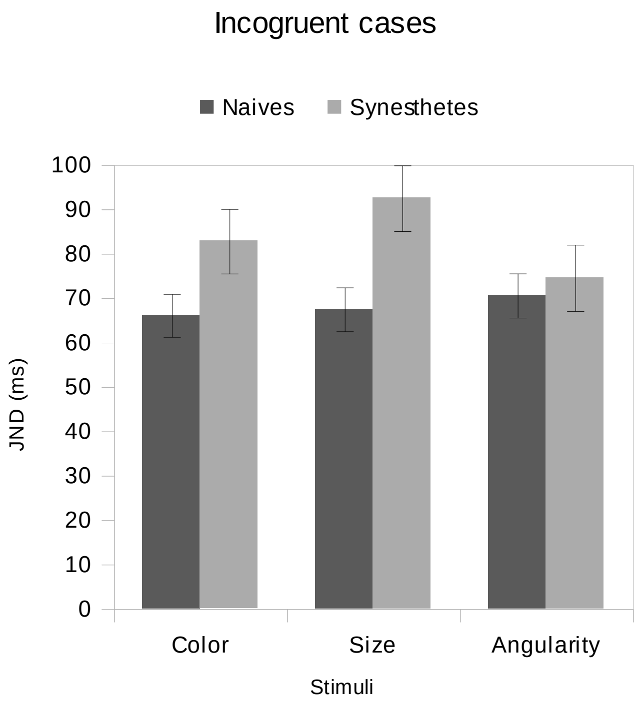

Figure 6: Synesthetes have higher JNDs in incongruent cases of the 'color' and the 'size' stimulus pairs. We observed the same trend for the ‘angularity’ pair.

Interestingly, we observed a significant effect of the interaction of Group*Congruency, for the incongruent cases [F(1,50)=5.746, p=0.02], with synesthetes (M=83) having higher JNDs than naïves (M=68; see Figure 5). Post-hoc analysis of this interaction revealed a significant difference in the incongruent case for the 'size' stimulus pairs [F=(1,50=7.891, p=0.007)], a marginal difference in the incongruent case for the 'color' stimulus pairs [F(1,50)=3.655, p=0.062], but no difference for the 'angularity' stimulus pairs [F(1,50)=0.197 p=0.659; see Figure 6].

2.4 Discussion

The results of the experiment reported here support previous findings of multisensory integration in crossmodal correspondences (Parise & Spence, 2009) using psychophysical tests. Specifically, we were able to demonstrate a 'unity effect' in naïves regarding the association of color, size, and angularity with pitched tones. Although further post hoc analyses on the JNDs of the congruent and incongruent case of each pair of stimuli separately wasn't statistically significant, we did observe the same trend in all three pairs of stimuli (see Figure 7). In our experiment, stimuli

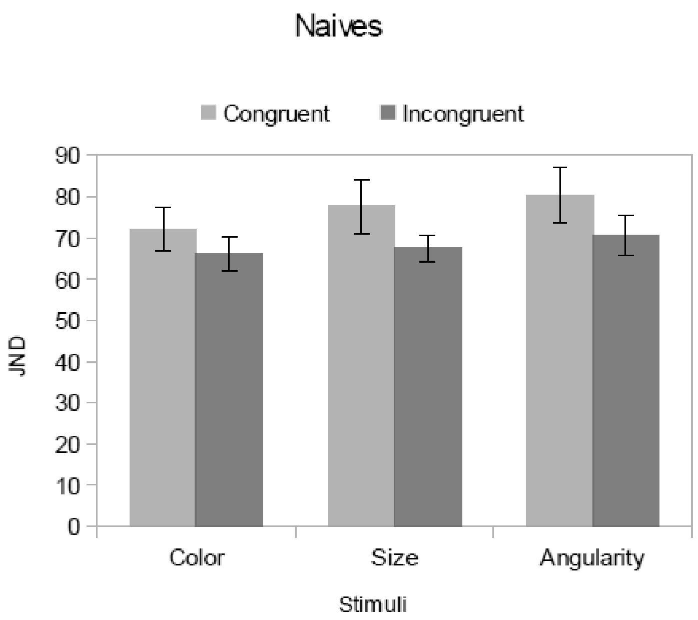

Figure 7: Naïve participants had grater difficulty in defining the temporal order of the auditory and the visual stimuli for every pair when the visual stimulus matched the auditory one, leading to a higher JND value for the congruent cases.

were presented randomly leading to a more robust experimental design. However, crossmodal correspondences depend largely on relative stimulus values (Marks, 1989; Parise & Spence, 2009), which could have affected the strength of the 'unity effect' experienced by the participants in our experiment. For instance, a large circle is perceived as “large” when compared to a smaller one, but since in our design the order of each stimulus was random, the comparison of the size of the stimuli was weaker.

In regards to the synesthetic data, we were able to collect data from 8 synesthetes, a relatively small sample, resulted to no significant differences; however it allowed us to compare multisensory integration between naïves and synesthetes. Synesthetes had higher JNDs both in congruent and in incongruent cases than naïves. These data support the idea that audiovisual synesthetes are better audiovisual integrators than naïves (Brang et al., 2012) irrespective of congruency. Brang, Williams, and Ramachandran (2012) showed data suggesting that synesthetes are better multisensory integrators than naïves using a paradigm based on the double flash illusion and a task of intersensory facilitation of reaction times for simple auditory, visual and audiovisual stimuli. In the congruent cases, although the stimuli are simple, there are additional processes involved, either cognitive or/and perceptual, that lead to the 'unity effect' and perhaps make the enhanced audiovisual integration more difficult to be observed.

3. Synesthetic Associations

3.1 Introduction

In grapheme-color synesthesia, the visual or auditory presence of a letter induces the experience of a specific color, while in number-color synesthesia the synesthetic color is co-experienced with the presence of a number. In the previous experiment we investigated the link between crossmodal correspondences and multisensory integration, but whether multisensory integration is involved in synesthesia remains an open question. In an attempt to answer this question, we will test synesthetes in their idiosyncratic audiovisual pairs of grapheme/number-color synesthesia with an unspeeded TOJ task. We will present audiovisual pairs in synesthetically congruent and incongruent and observe a potential difference in the JND values. If multisensory integration is involved in synesthesia, synesthetes should find it harder to judge which modality came first when presented with their idiosyncratic congruent pairs than in incongruent ones.

3.2 Methods

3.2.a Participants

Five grapheme-color and/or number color associator synesthetes were tested (M=27 years of age). All of them reported normal or corrected to normal vision and audition. Synesthesia has been confirmed using the Eagleman battery (Eagleman et al., 2007) and the associator-projector questionnaire of (Skelton et al., 2009). All participants signed a consent form and received monetary reward for their participation.

3.2.b *Apparatus & Materials*

Participants were presented with auditory and visual stimuli. Visual stimuli were presented on a 13.3-inch computer screen (refresh rate 60Hz) against a black background. Auditory stimuli were presented via 2 speakers placed at 3 cm to either side of the computer screen. Visual stimuli used as inducers for synesthesia consisted of 4 white letters of the latin alphabet for grapheme-color synesthetes and of 4 white numbers for the number-color synesthetes. We choose graphemes and numbers that scored low at the Eagleman test for each synesthete (a low score at the Eagleman battery indicates a more consistent association). Visual stimuli used as concurrents consisted of colored squares (230x230 pixels) at the exact tone of color that each synesthete associates with the chosen inducers. Auditory stimuli consisted of the letters, colors, and numbers selected for each synesthete pronounced by a female voice.

3.2.c *Procedure*

Synesthetes were asked to confirm whether the color stimuli were the same as their synesthetic color for the letters/numbers. Changes were made when indicated. Participants performed a TOJ task regarding which stimulus was presented first (i.e., visual or auditory) by pressing one of the response keys (i.e., “V” for visual, “A” for audio). Experimental duration was 50 minutes approximately.

3.2.d Design

All stimuli were presented for 300 ms in various SOAs (±250, ±150, ±100, ±80, 0 ms). Negative values indicate that the auditory stimulus was presented first and positive values that the visual stimulus was leading. The SOAs varied on a trial-by-trial basis using the method of constant stimuli. The experiment was separated in two parts regarding the modality of the synesthetic inducer (see figure 8). For the visual modality, the visual stimuli consisted of the letter/number stimuli and the auditory of the relevant colors, while for the auditory modality, the auditory stimuli consisted of the concurrent color and the visual of the letter/number stimuli. The experiment consisted of 6 blocks of 72 trials for each modality. The sequence of the modality parts were randomized across participants.

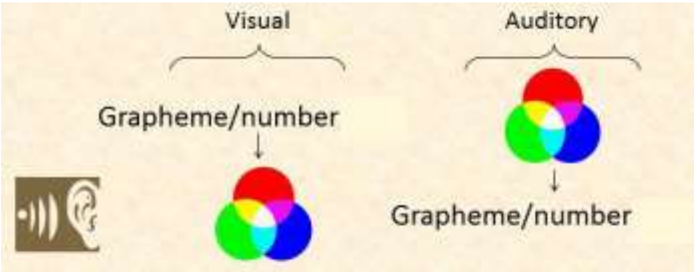

Figure 8: For the visual modality, the visual stimuli consisted of the grapheme/number stimuli and the auditory stimuli of the idiosyncratically associated colors, while for the auditory modality, the auditory stimuli consisted of the concurrent color and the visual one of the letter/number stimuli.

3.3 Results

One-way ANOVA with the factor of 'Modality' (auditory or visual) didn't reveal any difference between modalities so we were able to collapse the data regarding the modality of the inducer. Repeated measures ANOVA with the factor of 'Congruency' ('Inducer-concurrent congruent', 'Inducer-concurrent incongruent') revealed no difference between the JNDs of congruent (M=99 ms) and incongruent cases (M=85 ms) [$F(1,19)=0.6$, $P=0.45$; see Figure 9]. Results should be interpreted with caution due to the small number of synesthetes recruited so far.

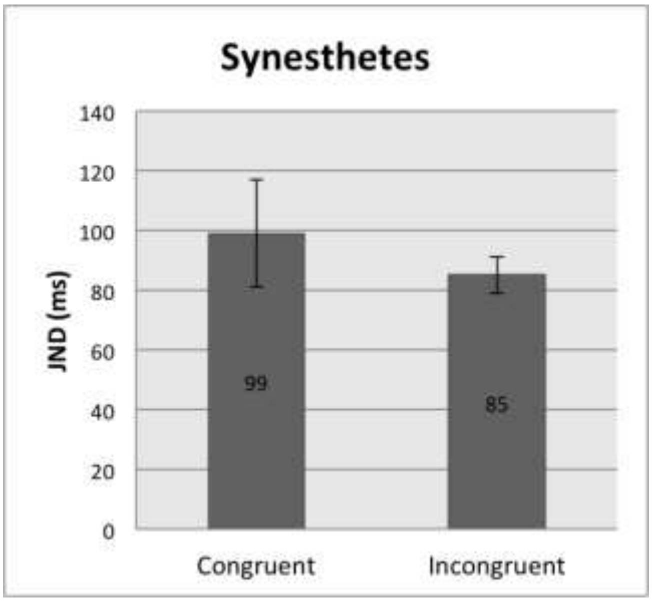

Figure 9: Synesthetes performed equally well when presented with congruent and with incongruent pairs of stimuli. This figure shows the mean JND value for each case.

3.4 Discussion

Synesthetes experience vivid, reliable color upon viewing achromatic alphanumeric characters. There is evidence that synesthetic color experiences are as perceptually real as actual colors are for non-synesthetes, though the interaction –if any- between the real and the synesthetic color in the synesthetes' mind is not yet clear (Kim, Blake, & Palmeri, 2006). Synesthetic colors can interfere with the naming of the real color of the presented grapheme, as shown with a synesthetic Stroop task (Dixon et al., 2000; Mattingley et al., 2001). Moreover, the background color can hamper the identification of a letter whose color matches the background (Smilek et al., 2001). The evidence for the interaction of the perceptual experience of the real and the synesthetic color highlight the distinct nature of these colors in the synesthetic mind (Sagiv & Robertson, 2005). For example, a number colored in red is seen as red with an additional synesthetic color superimposed on it. Van Leeuwen, Petersson, and Hagoort (2010) proposed that synesthetic color experiences are mediated by higher-order visual pathways that lie beyond the scope of classical, ventral-occipital visual areas of real color perception.

In our experiment, we attempt to measure integration between a visual and an auditory stimulus. In our synesthetic congruent cases, in order to capture the synesthetic color, we presented a real color that matches the synesthetic one. Perhaps, the absence of an effect of integration can be due to the native characteristic of the TOJ task to measure the JND values of the actual stimuli presented. Thus, since the actual and synesthetic colors are perceptually distinct, we examined the integration between an actual color that matches the synesthetic one and one that does not necessarily matches the experience. So, the nature of our task in combination with our small number of synesthetes can make the effect of integration difficult to emerge and possibly explain our null result.

Another aspect of synesthesia that might have tampered with our task is the unidirectionality of synesthetic associations. For example, in grapheme-color synesthesia, an achromatic letter elicits a specific color, but this color does not cause the experience of the letter. However, there is evidence that the presence of the synesthetic concurrent might trigger the synesthetic inducer (Gevers, Imbo, Cohen, Fias, & Hartsuike, 2010; Weiss, Kalckert, & Fink, 2009). In our experimental design, the modality of the inducer precedes the modality of the concurrent in half of the trials and follows it for the rest trials. We have to examine whether this asymmetry in synesthesia has any impact on our symmetric distribution of stimuli. We are going to further investigate integration in synesthetic associations taking into account all aspects of synesthesia.

4. Synesthetic Associations in Naïves

4.1 Introduction

Despite the idiosyncratic way of the synesthetic associations, certain grapheme-color pairs occur with higher frequency than other among synesthetes (Rich, Bradshaw, & Mattingley, 2005; Simner et al., 2005; see Figure 10).

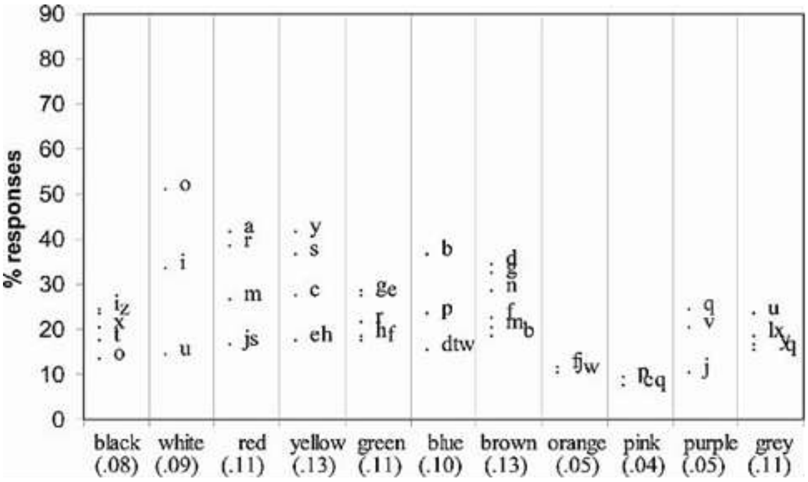

Figure 10: Significant letter–color combinations for synesthetes. Figure shows percentage of responses for each (significant) letter that were of a given color, with chance probabilities shown on the x-axis (Simner et al., 2005).

It has also been shown that naïves associate certain colors to certain letters frequently, when they are explicitly asked to do so (Simner et al., 2005; see Figure 11). We picked 4 grapheme-color pairs that are common in both synesthetes and naïves in order to investigate whether these combinations are stimulus driven or due to semantic correlation. If stimulus driven factors exist we should observe a "unity effect" for our grapheme-color pairs. Participants were presented with these grapheme-color pairs in congruent and in incongruent format and performed an unspeeded TOJ task. A "unity effect" in these common pairs should be interpreted as an indication of integration suggesting that there are certain factors in these stimuli, common to synesthetes and naïves that lead to their association. An absence of an effect would suggest that other, not stimulus driven, factors cause these association.

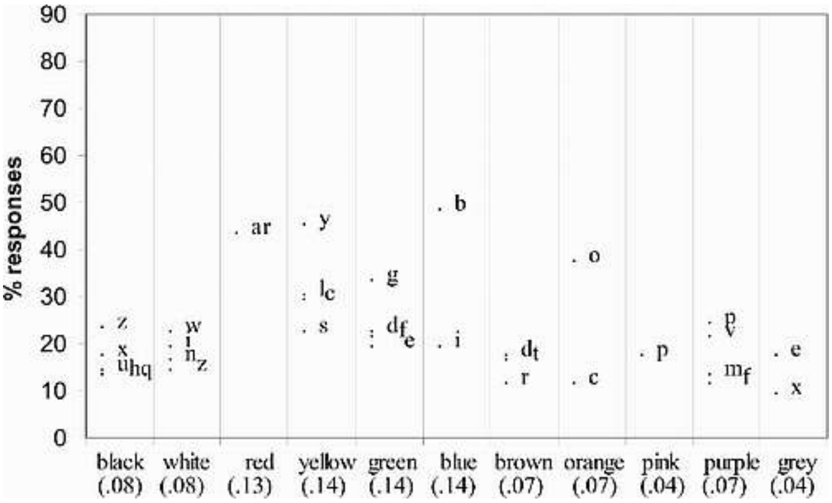

Figure 11: Significant letter–color combinations for naïves when asked explicitly to associate a color to a letter. Figure percentage of responses for each (significant) letter that were of a given color, with chance probabilities shown on the x-axis (Simner et al., 2005).

4.2 Methods

4.2.a Participants

Ten native English speakers (M=24) with no synesthesia reported were tested. Participants gave their written consent and received monetary reward.

4.2.b *Apparatus & Materials*

Participants were presented with auditory and visual stimuli. Visual stimuli were presented on a 13.3-inch computer screen (refresh rate 60Hz) against a black background. Auditory stimuli were presented via 2 speakers placed at 3 cm to either side of the screen. Visual stimuli consisted of

4 white capital letters of the latin alphabet (A, O, C, and G) and squares (230x230 pixels) colored in red, white, yellow, and brown. Auditory stimuli consisted of the letters and the colors pronounced by a female voice. We chose letters that naïves usually associate with certain colors (Rich et al., 2005). These letters are associated frequently with the same colors by grapheme-color synesthetes as well (Rich et al., 2005).

4.2.c *Procedure*

Participants received a fifteen-minutes training session on the letter-color associations and performed a TOJ task regarding which stimulus was presented first (i.e., visual or auditory). Experimental duration was 60 minutes approximately.

4.2.d Design

The design of the TOJ task was the same as in the second experiment. Instead for the inducer-concurrent synesthetic pair, we used the commonly associated colors to the 4 letters of the experiment. These associations formed the congruent cases. All participants were presented with the same letters/colors. For the training session participants received a card on which the letter-color associations were written and colored with the tone of the color mentioned and used in the experiment (e.g., 'A is RED', 'O is WHITE', 'C is YELLOW', 'G is BROWN'). Participants were asked to memorize those associations for as long as they needed. When they were ready, they moved on to the computer test. They were presented with each association for 3 seconds and asked to read out-loud. Every association was presented 3 times. Afterwards, the trials of the experimental session were presented but stimuli were in synchrony and participants were asked whether the pair of the visual and auditory stimuli matched the associations they have learned. Feedback was following every answer. The training test consisted of 2 blocks of X trials each. The card was taken back after the first block. If participants did more than 5 errors at the second block, they were asked to repeat the last block.

4.3 Results

Repeated measures ANOVA with the factors of 'Modality' (visual and auditory) and ‘Congruency’ (congruent and incongruent) was performed. Analysis revealed no main effect of ‘Modality’ [F(1,39)=0.35, P=0.5] and no main effect of ‘Congruency’ [F(1,39)=0.78, P=0.38]. Mean value for the congruent pairs was 81 ms while for the incongruent 85 ms (see Figure 12).

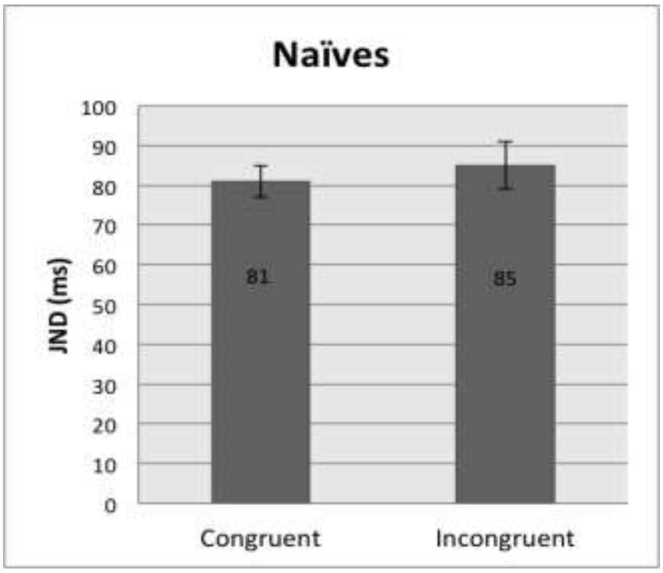

Figure 12: Participants were asked to make TOJs regarding which modality came first for letter-color combinations that has been shown to occur with high frequency among synesthetes and naïves (Simner et al., 2005) and for not so commonly associated combinations. Our experiment didn’t reveal any difference on participants’ ability to make the TOJ task due to the congruency factor.

4.4 Discussion

Our data do not reveal any differences between the associations that are typically matched and the unmatched ones. Presumably, our null result might be an indication towards the hypothesis that these common letter-color associations are due to semantic correlations (Simner, 2007). Further research is required to understand the nature of these associations in order to define the mechanisms that lead in the choice of certain colors by synesthetes and naïves. Another interesting question concerns the commonalities in the origin of these associations for both groups. Are they due to a shared mechanism between synesthetes and naïves or do they arise for different reasons in each group?

5. Conclusion

In this thesis, we explored the field of multisensory perception by focusing on the auditory and visual modalities and their interaction. We approach the perception of audiovisual stimuli through the paradigms and theory of crossmodal correspondences. Crossmodal correspondences' experimental designs mainly use simple audio and visual stimuli to investigate how one perceives an audiovisual percept as an integrated unique stimulus. Multisensory integration plays a crucial role in defining the factors that determine whether two stimuli from different modalities will be perceived as one multimodal stimulus or two separate unimodal stimuli (i.e., unity effect).

Integration is our primary focus and we investigated whether another multisensory phenomenon, synesthesia -in which certain stimuli elicit the automatic perception of a typically unrelated second stimulus- can help us highlight aspects of integration. So, we used both synesthetes and naïves as participants. This combination of multisensory phenomena can also be enlightening for synesthesia itself, since we use experimental tasks used typically for the study of multisensory integration for the first time on synesthetes. These tasks allowed us to explore integration in synesthesia using an objective task and, thus, allowing the study of the possible links between synesthetic associations of synesthetes and similar associations of naïves.

In our first experiment, we tested both synesthetes and naïves in crossmodal audiovisual correspondences concerning color, size, and angularity. We were able to reproduce previous results reporting a unity effect between audio and visual stimuli. Although for synesthetes our results didn't reach significance, we did observe the same trend. Regarding group differences, synesthetes found more difficult to make the TOJs between the visual and audio modality than naïves. The idea that synesthetes are better integrators has gained some research interest (Brang et al., 2012), but needs further experimentation.

Our initial attempt to investigate integration in synesthesia led to our second experiment in which our stimuli consisted of synesthetic idiosyncratic associations. Each synesthete was presented with pairs of stimuli based on his/her unique associations, which were either congruent or incongruent. As in the previous experiment, we used an unspeeded TOJ task. Our null result for the congruency factor in synesthetes could be due to lack of bidirectionality in synesthetic associations, since a letter might elicit a specific color but the color doesn't have the same effect. So, it could be argued that for the pairs of audiovisual congruent pairs were the modality of the concurrent was presented first, shouldn't be counted as "congruent cases" in our analysis. We need to further investigate whether the order of the synesthetic inducer and synesthetic concurrent plays a role in our task. Another notion we should further investigate is the individual differences of synesthetes. We used associator synesthetes, either grapheme-color or number-color. Associators experience the synesthetic concurrent in their "mind's eye". Our TOJ task measures integration of the inducer with the physical figure of the concurrent. Whether and how the presentation of a stimulus in "the mind's eye" instead of an externally displayed one tampers with audiovisual processing of the synesthetic pair should be explored in a future study. Furthermore, we should investigate how the virtual position of the inducer interacts with an external stimulus in the case of projector synesthetes where the synesthetic color is superimposed over the synesthetic inducer, as individual differences in synesthesia play an important role on experimental results (Rouw & Scholte, 2010).

Although synesthetes are usually treated as a separate experimental group, there are certain characteristics of synesthetes that are observed on naïves too. We tried to explore whether these behavioral commonalities have shared perceptual mechanisms. So, we tested naïves on frequently associated letters and colors, using a TOJ task. Our data showed that naïves find it equally difficult to make the TOJs for frequent and for random associations. Our finding leans toward the hypothesis that choices of non-synesthetes are influenced by order of elicitation, and by exemplar typicality from the semantic class of colors, while synesthetes tend to associate higher frequency graphemes with higher frequency color terms (Simner et al., 2005). We should further investigate the semantic and linguistic determinants of grapheme-color associations to further understand the developmental course of synesthesia and the acquisition of these associations.

In conclusion, the current thesis explored explore whether we can demonstrate a link between multisensory integration and the phenomenon of synesthesia. Such a link could provide useful information and a special experimental tool for multisensory perception, since synesthetes have explicit knowledge of their associations.

References

Bedford, F. (2001). Towards a general law of numerical/object identity. *Current Psychology of Cognition*, **20**, 113–176.

Beeli, G., Esslen, M., & Jäncke, L. (2008). Time Course of Neural Activity Correlated with Colored-Hearing Synesthesia. *Cerebral Cortex*, **18**, 379–385.

Bertelson, P., Frissen, I., Vroomen, J., & de Gelder, B. (2006). The aftereffects of ventriloquism: patterns of spatial generalization. *Perception & psychophysics*, **68**, 428–436.

Brang, D., Hubbard, E. M., Coulson, S., Huang, M., & Ramachandran, V. S. (2010). Magnetoencephalography reveals early activation of V4 in grapheme-color synesthesia. *NeuroImage*, **53**, 268–274.

Brang, D., Williams, L. E., & Ramachandran, V. S. (2012). Grapheme-color synesthetes show enhanced crossmodal processing between auditory and visual modalities. *Cortex*, **48**, 630–637.

Chen, Y.-C., & Spence, C. (2010). When hearing the bark helps to identify the dog: Semantically-congruent sounds modulate the identification of masked pictures. *Cognition*, **114**, 389–404.

Dixon M.J., Smilek D., Cudahy C. & Merikle P.M. (2000). Five plus two equals yellow. *Nature*, **406: 365**, 2000.

Dixon, M. J., & Smilek, D. (2005). The importance of individual differences in grapheme-color synesthesia. *Neuron*, **45**, 821–823.

Eagleman, D. M., Kagan, A. D., Nelson, S. S., Sagaram, D., & Sarma, A. K. (2007). A standardized test battery for the study of synesthesia. *Journal of Neuroscience Methods*, **159**, 139–145.

Ernst, M. O. (2006). A Bayesian view on multimodal cue integration. In G. Knoblich, I. Thornton, M. Grosejan, & M. Shiffrar (Eds.), *Perception of the human body from the inside out* (pp. 105–131). New York: Oxford University Press.

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of*

Vision, **7(5)**, 1–14.

Gevers, W., Imbo, I., Cohen Kadosh, R., Fias, W., & Hartsuiker, R. J. (2010). Bidirectionality in synesthesia: Evidence from a multiplication verification task. *Experimental Psychology*, **57(3),** 178–184.

Jones, J. A., & Jarick, M. (2006). Multisensory integration of speech signals: the relationship between space and time. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, **174**, 588–594.

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

Keetels, M., & Vroomen, J. (2011). No effect of synesthetic congruency on temporal ventriloquism. *Attention, perception & psychophysics*, **73**, 209–218.

Köhler, W. (1947). Gestalt psychology: An introduction to new concepts in modern psychology. New York: Liveright.

Ludwig, V. U., Adachi, I., & Matsuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans. *Proceedings of the National Academy of Sciences of the United States of America*. **108**, 20661-20665.

Mattingley, J.B., Rich, A.N., Yeland, G., & Bradshaw, J.L. (2006). Unconscious priming eliminates automatic binding of colour and alphanumeric form in synesthesia. *Nature*, **410**: 580-582.

Marks, L. E. (1989). On cross-modal similarity: the perceptual structure of pitch, loudness, and brightness. *Journal of experimental psychology. Human perception and performance*, **15**, 586–602.

Neufeld, J., Sinke, C., Zedler, M., Emrich, H. M., & Szycik, G. R. (2012). Reduced audio–visual integration in synaesthetes indicated by the double-flash illusion. *Brain Research*, **1473**, 78–86.

Parise, C., & Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neuroscience Letters*, **442**, 257–261.

Parise, C., & Spence, C. (2009). "When birds of a feather flock together": Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS ONE*, **4**, e5664.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies*, **8**, 3–34.

Ramachandran, V. S., & Hubbard, E. M. (2003). Hearing colors, tasting shapes. *Scientific American*, **288**, 52–59.

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.

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience*, **10**, 792–797.

Rouw, R., & Scholte, H. S. (2010). Neural basis of individual differences in synesthetic experiences. *The Journal of Neuroscience*, **30**(18), 6205–6213.

Sagiv, N., & Robertson L. C. (2005). Synesthesia and the binding problem. In L.C. Robertson & N. Sagiv (Eds.), *Synesthesia: Perspectives from Cognitive Neuroscience* (pp.90-107). New York: Oxford University Press.

Sagiv, N., & Ward, J. (2006). Cross-modal interactions: Lessons from synesthesia. In S. Martinez-Conde, S.L. Macknik, L.M. Martinez, J. Alonso, & P.U. Tse, (Eds), *Progress in Brain Research*, (pp. 263-275). London: Elsevier Science.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, **12**, 225–239.

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**, 1069–1085.

Simner, J. (2007). Beyond perception: Synaesthesia as a psycholinguistic phenomenon. *Trends in*

Cognitive Sciences, **11**(1), 23–29.

Skelton, R., Ludwig, C., & Mohr, C. (2009). A novel, illustrated questionnaire to distinguish projector and associator synaesthetes. *Cortex*, **45**, 721–729.

Smilek, D., Dixon, M.J., Cudahy, C., & Merikle, P.M. (2001). Synesthetic photisms influence visual perception. Journal of Cognitive Neuroscience, 13: 930-936.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, **73**, 971–995.

Stein, B. E., & Meredith, M. A. (1993). The Merging of the Senses. Cambridge: MIT Press.

van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, **43**, 271–282.

van Leeuwen, T. M., Petersson, K. M., & Hagoort, P. (2010). Synaesthetic colour in the brain: Beyond colour areas. A functional magnetic resonance imaging study of synaesthetes and matched controls. *PLoS ONE*, **5**(8), e12074.

Vatakis, A., & Spence, C. (2007). Crossmodal binding: Evaluating the "unity assumption" using audiovisual speech stimuli. *Attention, Perception, & Psychophysics*, **69**, 744–756.

Vroomen, J., & Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of experimental psychology: Human perception and performance*, **26**, 1583.

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

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, **21,** 21–25.

Weiss, P. H., Kalckert, A., & Fink, G.R. (2009). Priming letters by colors: Evidence for the bidirectionality of grapheme-color synesthesia. *Journal of Cognitive Neuroscience*, **21(10),** 2019–2026.

Welch, R. B. (1999a). Meaning, attention, and the "unity assumption" in the intersensory bias of spatial and temporal perceptions. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Advances in Psychology* (pp. 371–387). Amsterdam: Elsevier.

Welch, R. B. (1999b). The advantages and limitations of the psychophysical staircases procedure in the study of intersensory bias: Commentary on Bertelson. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Advances in Psychology* (pp. 363–369). Amsterdam: Elsevier.

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