

TESIS DOCTORAL

**PROCESAMIENTO MULTISENSORIAL Y ATENCIÓN: EFECTOS DE LA
PROPORCIÓN DE CONGRUENCIA**

DOCTORANDA

BEATRIZ RODRÍGUEZ SARMIENTO

DIRECTOR

DANIEL SANABRIA LUCENA

DEPARTAMENTO DE PSICOLOGÍA EXPERIMENTAL



UNIVERSIDAD DE GRANADA

OCTUBRE 2014

Editor: Editorial de la Universidad de Granada
Autor: Beatriz Rodríguez Sarmiento
D.L.: GR 2322-2014
ISBN: 978-84-9083-369-8



UNIVERSIDAD DE GRANADA

Universidad de Granada
Departamento de Psicología Experimental
Campus Universitario de Cartuja, s/n
Telf. +34-958243763 –Fax: +34 -958246239
18071 – Granada (España)

Procesamiento multisensorial y atención: efectos de la proporción de congruencia

Tesis Doctoral presentada por **Beatriz Rodríguez Sarmiento** en el *Departamento de Psicología Experimental* para aspirar al grado de Doctora en Psicología, en el Programa de *Doctorado en Psicología* de la Universidad de Granada. En este trabajo se han respetado las pautas que establece la normativa de la Universidad de Granada para la obtención del título de Doctorado Internacional.

La tesis ha sido realizada bajo la dirección del profesor Daniel Sanabria Lucena, quien avala la calidad de la misma, así como la formación de la doctoranda para aspirar al grado de doctor.

La doctoranda Beatriz Rodríguez Sarmiento y el director de la tesis Daniel Sanabria Lucena garantizan, al firmar esta tesis doctoral, que el trabajo ha sido realizado por la doctoranda bajo la dirección del director de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Firmado en Granada, a 2 de Septiembre de 2014

La doctoranda:

El director de tesis:

Fdo.: Beatriz Rodríguez Sarmiento

Fdo.: Daniel Sanabria Lucena

Los estudios presentados en esta tesis han sido financiados a través de una beca del Plan Propio de la Universidad de Granada, dentro del programa de *Formación de Investigadores* y una beca del Ministerio de Educación y Ciencia (AP2008-03662), dentro del programa nacional de *Formación del Profesorado Universitario* (FPU).

Al héroe epónimo

Esta tesis ha supuesto muchos días de trabajo académico, pero tras ella se acumulan también muchos días de crecimiento y aprendizaje personal. Ambos logros no hubiesen sido posibles sin la presencia de todas las personas con las que he tenido la suerte de contar.

Agradecer a Dani su cercanía e infinita paciencia a la hora de guiarme en este trayecto, tanto a través de su conocimiento como de sus frases de aliento en los momentos en que una cree que la única opción es tirar la toalla.

Gracias a mi familia por apoyarme y animarme en la consecución de mis objetivos. Por estar siempre presentes y disponibles. Por acostumbrarse a la distancia y contar los días para el próximo reencuentro con una sonrisa en los labios.

No imagino mis últimos años sin las actuaciones, las fiestas del colegio, los partidos y campeonatos, y sobre todo los domingos de anécdotas acompañados de papas a lo pobre. Gracias por hacerme parte de la familia desde el primer día.

Qué importantes han sido también las noches de japonés, los paseos por Granada a la hora y el momento que fuese necesario, las comidas improvisadas y las cenas organizadas, las visitas a la Alhambra, las limonadas con hierbabuena, los espectáculos de flamenco, los viajes de fin de semana, las compañeras de piso y las vecinas de abajo, el *toc toc* de la puerta del despacho previo a una buena charla o un buen té, las llamadas de teléfono que hacen desaparecer los kilómetros, los reencuentros navideños, estivales y alguno que otro fuera de temporada...

Y, fundamentalmente, qué importante ha sido recordar que cualquier objetivo se puede lograr cuando tienes a alguien que te da la mano en el camino.

ÍNDICE DE CONTENIDOS

ÍNDICE DE CONTENIDOS	9
ÍNDICE DE FIGURAS Y TABLAS	15
CAPÍTULO 1. INTRODUCCIÓN.....	19
<u>1. Procesamiento multisensorial: Una visión general</u>	22
1.1. Primeros estudios: definiendo el fenómeno.....	23
1.2. Factores determinantes: físicos y estructurales.....	24
1.3. Factores determinantes: cognitivos	27
1.4. Paradigmas para el estudio del procesamiento multisensorial: Conflicto.....	28
1.5. Otras aproximaciones: Nivel de procesamiento	30
1.6. Mecanismos neurales.....	33
<u>2. Atención y procesamiento multisensorial</u>	37
2.1. Atención.....	38
2.2. Primeras aproximaciones	41
2.3. Perspectiva actual.....	42
<u>3. Congruencia: Manipulando las proporciones</u>	47
3.1. Efecto de congruencia: índice de procesamiento multisensorial.....	48
3.2. Efecto de proporción de congruencia específico del contexto	50
3.3. Paradigma CSPC y atención	53
CAPÍTULO 2. MOTIVACIÓN Y OBJETIVOS	55
CAPÍTULO 3. SERIE EXPERIMENTAL I. AUDIOVISUAL INTERACTIONS DEPEND ON CONTEXT OF CONGRUENCE	65
<u>1. Abstract</u>	67
<u>2. Introduction</u>	67
<u>3. Experiment 1</u>	71
3.1. Method.....	71
3.1.1. <i>Participants</i>	71
3.1.2. <i>Apparatus and stimuli</i>	72
3.1.3. <i>Procedure</i>	72
3.2. Results	73
3.3. Discussion.....	75
<u>4. Experiment 2</u>	76
4.1. Method.....	76
4.1.1. <i>Participants</i>	76
4.1.2. <i>Apparatus, stimuli and procedure</i>	76

CONTENIDOS

4.2. Results	77
4.2.1. Accuracy measures	77
4.2.2. Questionnaire measures	79
4.3. Discussion.....	81
<u>5. Experiment 3</u>	82
5.1. Method	82
5.1.1. Participants	82
5.1.2. Apparatus, stimuli and procedure	82
5.2. Results	83
5.2.1. Accuracy measures	83
5.2.2. Questionnaire measures	84
5.3. Discussion.....	84
<u>6. Experiment 4</u>	85
6.1. Method	85
6.1.1. Participants	85
6.1.2. Apparatus, stimuli and procedure	85
6.2. Results	85
6.2.1. Accuracy measures	85
6.2.2. Questionnaire measures	87
6.3. Discussion.....	88
<u>7. General Discussion</u>	90
CAPÍTULO 4. SERIE EXPERIMENTAL II. ELECTROPHYSIOLOGICAL EVIDENCE OF AUDIOVISUAL INTERACTIONS UNDER DIFFERENT CONTEXTS OF STIMULUS CONGRUENCE.....	95
<u>1. Abstract</u>	97
<u>2. Introduction</u>	98
<u>3. Method</u>	101
3.1. Participants	101
3.2. Apparatus and stimuli.	101
3.3. Procedure	102
3.4. EEG acquisition	104
3.5. ERP analyses	105
3.6. Source Estimations	107
<u>4. Results</u>	108
4.1. Oddball Task.....	109

4.1.1. <i>Behavioural data</i>	109
4.1.2. <i>Surface ERP data</i>	109
4.1.3. <i>Source estimations</i>	111
4.2. Visual Temporal Discrimination Task.....	112
4.2.1. <i>Behavioural data</i>	112
4.2.2. <i>Questionnaire measures</i>	114
4.2.3. <i>Surface ERP data</i>	114
4.2.4. <i>Source estimations</i>	118
5. <u>Discussion</u>	120
CAPÍTULO 5. DISCUSIÓN.....	125
1. <u>Aportaciones metodológicas</u>	128
2. <u>Resumen de los resultados principales</u>	130
3. <u>Integración audiovisual</u>	134
4. <u>Integración audiovisual y mecanismos de selección de información</u>	138
REFERENCIAS BIBLIOGRÁFICAS	145
ANEXO I. EXPERIMENT 5.....	163
1. <u>Experiment 5</u>	165
1.1. Method	165
1.1.1. <i>Participants</i>	165
1.1.2. <i>Apparatus, stimuli and procedure</i>	165
1.2. Results	166
2.1.1. <i>Behavioural data</i>	166
2.1.1. <i>Questionnaire measures</i>	168
2.1.2. <i>Surface ERP data</i>	168
2.1.3. <i>Source estimations</i>	170
2. <u>Technical issues</u>	171

ÍNDICE DE FIGURAS Y TABLAS

Figura 1. Áreas implicadas en el procesamiento multisensorial en el cerebro humano..... 36

Figure 2. Schematic illustration of the set up used in Experiments 1-3 (A) and Experiment 4 (B)..... 73

Figure 3. Percentage of correct responses for Unimodal and Crossmodal (congruent and incongruent) trials in Experiment 1..... 74

Figure 4. Percentage of correct responses as a function of Congruence, Proportion congruent and Block in Experiment 2..... 78

Figure 5. Percentage of correct responses as a function of Congruence, Proportion congruent and Context difference in Experiment 2 80

Figure 6. Percentage of correct responses as a function of Congruence, Proportion congruent and Block in Experiment 3..... 84

Figure 7. Percentage of correct responses as a function of Congruence, Proportion congruent and Block in Experiment 4..... 87

Figure 8. Percentage of correct responses as a function of Congruence, Proportion congruent and Context difference in Experiment 4 88

Figure 9. Schematic illustration of the set up used in the Experiment 104

Figure 10. Group-average GFP waveforms and results of t-tests on the GFP as a function of time..... 109

Figure 11. a. Template maps from topographic cluster analysis. **b.** Results of the spatial correlation fitting over 115-154 ms period in the single-subject data from each condition 111

Figure 12. Results of the source estimations from the 115-154 ms post-stimulus onset time window..... 112

Figure 13. Percentage of correct responses as a function of Context transition, Context and Congruence..... 114

Figure 14. Group-average GFP waveforms and results of the ANOVA on the GFP as a function of time 115

Figure 15. a. Template maps from topographic cluster analysis. **b.** Results of the spatial correlation fitting in the single-subject data from each condition..... 117

Figure 16. Group-average GFP waveforms and results of the ANOVA on the GFP as a function of time 118

Figure 17. Results of the post-hoc comparisons between congruent and incongruent stimulation in the high proportion-congruent context over the 163-227 ms interval..... 119

Figure 18. Schematic illustration of the set up used in the Experiment. 166

Figure 19. Percentage of correct responses as a function of Context transition, Context and Congruence..... 168

Figure 20. Group-average GFP waveforms and results of the ANOVA on the GFP as a function of time 169

Figure 21. a. Template maps from topographic cluster analysis. **b.** Results of the spatial correlation fitting in the single-subject data from each condition..... 170

Figure 22. Results of source estimations ANOVA between the high and low proportion-congruent contexts over the 116-216 ms interval 171

Figure 23. Waveforms of Experiment 5 (**a**) and Experiment 6 (**b**)..... 172

Table 1. Mean correct responses and standard errors, for Experiments 2-4 78

Table 2. Percentage of participants that noticed or not any difference between the two contexts of congruence in Experiments 2-4 79

Table 3. Percentage of participants who considered feedback information as reliable or unreliable in Experiments 2-4 81

CAPÍTULO 1. INTRODUCCIÓN

Imaginemos por un momento que estamos sentados en un parque observando la escena. Hay mucho bullicio: niños que juegan, jóvenes que hablan, pájaros que revolotean, bicicletas que pasan... En este momento la información del ambiente que nos rodea llega a nosotros a través de distintos de nuestros sentidos, así, mediante la vista observamos a los niños, las bicicletas...; mediante el oído escuchamos la conversación de los jóvenes, el trino de los pájaros...; el tacto nos informa de las características de la superficie en la que estamos sentados (v.gr., un banco, la hierba...), de la brisa que provoca una bicicleta pasando a poca distancia... Toda esa estimulación que nos llega nos permite crear una imagen precisa de la realidad que nos rodea. Una realidad que, como acabamos de mostrar, es fundamentalmente multisensorial, requiriendo de todos los sentidos para construirse. Muy importante es caer en la cuenta de que esta percepción de la realidad requiere de la combinación de la información que las distintas modalidades sensoriales aportan. Volviendo a nuestro ejemplo, nos hacemos una imagen global de la escena cuando integramos la imagen de los jóvenes con el sonido de sus palabras y no con el trino de los pájaros, o cuando asociamos la imagen de la bicicleta, el sonido de la cadena y la brisa que deja con su movimiento, en lugar de asociarla con el sonido y movimiento de aire producido por el aleteo de los pájaros. El estudio de cómo el cerebro integra la información que llega desde distintas modalidades sensoriales, caracterizadas por receptores sensoriales independientes que dirigen la información a áreas cerebrales bien diferenciadas, es lo que se ha venido en denominar *procesamiento multisensorial* (véase Stein, 2012; Murray y Wallace, 2012; Stein y Meredith, 1993). Concretamente desde el ámbito de la neurociencia cognitiva, dentro del cual se enmarcan los estudios que conforman esta tesis, se pretenden estudiar los mecanismos neurales que subyacen al procesamiento multisensorial, así como sus manifestaciones conductuales.

1. Procesamiento multisensorial: Una visión general

El término *procesamiento multisensorial* se refiere a los procesos mediante los cuales la información que múltiples sentidos proporcionan sobre un mismo evento se combina en el cerebro para conformar la experiencia perceptiva. Estos procesos incluyen desde modulación de la información ofrecida por una modalidad sensorial por la información que ofrece otra modalidad, - por ejemplo la ventriloquia, donde se percibe la voz del ventrílocuo (i.e., modalidad auditiva) proviniendo de la localización espacial en la que se mueven los labios de una marioneta (i.e., modalidad visual) -, hasta la creación de un nuevo percepto distinto al que se derivaría de las modalidades sensoriales implicadas, por ejemplo el efecto McGurk (McGurk y MacDonald, 1976), en el que la combinación del sonido /ba/ con el movimiento labial /ga/ produce la percepción de un tercer sonido /da/. A las distintas formas de combinar la información se las identifica con distintos fenómenos, así hablaríamos de *facilitación* cuando la presencia de estimulación en una segunda modalidad mejora la percepción de la información en una primera modalidad; de *emparejamiento multisensorial* cuando se comparan las características que dos modalidades sensoriales comparten sobre un mismo evento; de *interacción multisensorial* cuando una modalidad sensorial está modulada por la presencia de otra pero que no da lugar a una percepción distinta a la que se derivaría de la percepción de cada uno de los componentes unisensoriales presentados; o de *integración multisensorial* cuando la percepción final sí implica una respuesta diferente a la mera combinación de ambos estímulos; etc. (véase Calvert, 2001; Stein, Burr, Constantinidis y cols., 2010 para una discusión terminológica más profunda).

Desde el punto de vista de la neurociencia cognitiva el estudio del procesamiento multisensorial ha mostrado los mecanismos responsables de dichos fenómenos a nivel tanto anatómico como funcional (v.gr., Calvert, Spence, y Stein, 2004; Spence y Driver, 2004; Pascual-Leone y Hamilton, 2001; Shimojo y Shams, 2001; Welch y Warren, 1980). Se ha observado, por ejemplo, la presencia de múltiples conexiones entre las áreas

unisensoriales (i.e., neuronas que responden a inputs de una única modalidad sensorial; v.gr, Falchier, Clavagnier, Barone, y Kennedy, 2002; Rockland y Ojima, 2003) así como de áreas especializadas en la integración (i.e., neuronas que responden a más de un tipo de información sensorial), cuya existencia ha sido demostrada desde el nivel de registro unicelular (véase Stein y Meredith, 1993) hasta el de técnicas de neuroimagen (véase Murray y Wallace, 2012).

1.1. Primeros estudios: definiendo el fenómeno

Las primeras investigaciones llevadas a cabo en el área del procesamiento multisensorial cubrían dos temáticas fundamentales, por un lado determinar a nivel neural un índice de integración multisensorial, y por otro identificar a nivel conductual la modulación que unas modalidades sensoriales ejercen sobre las otras.

A nivel neural los estudios que se centraron en el registro unicelular animal encontraron que la tasa de respuesta de las neuronas a estímulos multisensoriales era significativamente diferente a la combinación de las tasas de respuesta a cada uno de los estímulos presentados separadamente, tanto porque la tasa de respuesta a estímulos multisensoriales era mayor que la suma de las respuestas a los estímulos unisensoriales - *superaditividad* o *incremento* - (v.gr., Meredith y Stein, 1986), como porque la tasa de respuesta era menor de lo que se esperaría tomando las respuestas a los estímulos presentados separadamente - *subaditividad* o *depresión* - (v.gr., Meredith y Stein, 1983). Así pues, esta diferencia cualitativa en la respuesta a estímulos multisensoriales se ha tomado como índice de integración multisensorial no sólo a nivel unicelular sino también a nivel de registro no invasivo, como pueden ser la electroencefalografía (v.gr., Giard y Peronet, 1999; Talsma, Doty y Woldorff, 2007) o la resonancia magnética (v.gr., Calvert, Campbell y Brammer, 2000; Macaluso y Driver, 2005), e incluso a nivel conductual con modelos bayesianos que infieren las características de integración en el cerebro (véase

Alais y Burr, 2004 para información audiovisual; Ernst y Banks, 2002 para información viso-táctil).

Los estudios sobre el correlato comportamental del procesamiento multisensorial se centraban en la observación de fenómenos de facilitación y de competición entre estímulos. Por facilitación se entiende la mejora en la ejecución de alguna tarea debido a la presentación simultánea de información en más de una modalidad sensorial, por ejemplo, la disminución en el tiempo de reacción a la hora de detectar un estímulo visual cuando éste va acompañado de un estímulo auditivo que ofrece información redundante (i.e., efecto de redundancia sensorial [RSE; v.gr., Kinchla, 1974]), o incluso cuando el estímulo auditivo es irrelevante para la tarea (v.gr., Nickerson, 1973). Al hablar de competición entre estímulos nos referimos a como la información en una modalidad sensorial disminuye el peso de la información en otra modalidad sensorial. El clásico efecto Colavita (Colavita, 1974) es un ejemplo de competición donde, debido a la dominancia de la modalidad visual sobre la auditiva en el ser humano, los estímulos auditivos no suelen detectarse cuando se presentan en sincronía con los visuales, mientras sí se detectan correctamente cuando se presentan de forma aislada.

1.2. Factores determinantes: físicos y estructurales

El enfoque de los primeros estudios sobre procesamiento multisensorial se ha mantenido a lo largo de los años con una doble intención. Por un lado, tratar de afianzar los resultados obtenidos (veáse Sinnet, Soto-Faraco y Spence, 2007; Koppen y Spence, 2007a,b para estudios sobre el efecto Colavita; veáse Stein, 1998; Avillac, Hamed y Duhamel, 2007 para estudios sobre superaditividad y subaditividad neural), y por otro lado, tratar de encontrar nuevos efectos que ayudasen a entender como el cerebro procesa estímulos multisensoriales. Por ejemplo, Shams, Kamitani, y Shimojo (2000) mostraron como la percepción del número de estímulos visuales se ve modula por la presentación de estímulos auditivos: percibiéndose dos flashes cuando se presenta un

solo estímulo visual acompañado de dos beeps - *ilusión del doble flash* (véase Andersen, Tiippana y Sams, 2004 para ilusiones complementarias) -. Además, una vez había sido descrito el fenómeno de integración multisensorial a nivel neural y conductual, la investigación centró el foco de estudio en aquellos factores que se consideraban determinantes para, mediante su manipulación, conocer las propiedades inherentes a este tipo de procesamiento. Entre estos factores se encuentran los parámetros espaciales y temporales de la información sensorial, la estructura física y funcional de los órganos sensoriales o la calidad de información que llega a cada sentido.

Las características estructurales y funcionales de nuestros sentidos juegan un papel importante a la hora de determinar el peso de la información recibida y su consiguiente integración. Así, es bien sabido que la modalidad visual se caracteriza por dominar sobre el resto de la modalidades cuando a localización se refiere, debido a su mayor resolución en términos de procesamiento espacial (v.gr., Pick, Warren y Hay, 1969; Freides, 1974; Welch y Warren, 1980; Alais y Burr, 2004; Witten y Knudsen, 2005). Mientras que si nos movemos más en un dominio temporal (v.gr., duración), será la modalidad auditiva la que, por su mayor resolución en este tipo de procesamiento, tendrá más peso (v.gr., Shams y cols., 2000; Bertelson y Archersleben, 2003; Morein-Zamir, Soto-Faraco, y Kingstone, 2003; Getzman, 2007; Klink, Montijn y van Wezel, 2011). Estrechamente relacionada con la dominancia de cada modalidad sensorial está la calidad de la información recibida. Aunque la visión domine en el espacio, si la calidad de la señal visual es tenue, por ejemplo, mediante la manipulación de las propiedades que nos permiten tener una visión tridimensional de un objeto, otra modalidad sensorial podrá mostrar más peso a la hora de generar el percepto final (véase Ernst, Banks y Bühlhoff, 2000 para información viso-táctil; Alais y Burr, 2004 para información audiovisual). Del mismo modo, un sonido presentado muy cercano al nivel del umbral auditivo perderá su dominancia en el procesamiento temporal frente a estimulación visual (véase Andersen, Tiippana, y Sams, 2004). De modo que uno de los principios importantes que rige el procesamiento multisensorial es cuán

apropiada es la modalidad sensorial para el tipo de procesamiento a llevar a cabo (i.e., modality appropriateness hypothesis, v.gr., Battaglia, Jacobs y Aslin, 2003; Ernst y Bulthoff, 2004; Deneve y Pouget, 2004; Welch y Warren, 1980), así como cuán efectiva es la señal, puesto que cuanto mayor sea el ruido menor peso tendrá esa modalidad sensorial en la percepción final (i.e., inverse effectiveness a nivel neural, v.gr., Sumbly y Pollack, 1954; Perrault y cols., 2003; Stanford, Quessy y Stein, 2005; i.e., Maximum Likelihood Estimation model (MLE), v.gr., Ernst y Banks, 2002; Pouget, Deneve y Duhamel, 2002; Alais y Burr, 2004; Kersten, Mamassian y Yuille, 2004).

En cuanto a las propiedades físicas de los estímulos, las características temporales y espaciales han sido las más ampliamente estudiadas, demostrándose que la coincidencia espacial y/o temporal mejora el procesamiento multisensorial. Así, ya en 1975 Choe, Welch, Gilford y Juola encontraron una disminución del efecto del ventrílocuo cuando la presentación de ambos estímulos era asincrónica. Del mismo modo, el efecto del ventrílocuo disminuye cuando la distancia entre la fuente del sonido y la fuente de la estimulación visual aumenta (v.gr., Slutsky y Recanzone, 2001). Shams, Kamitani, y Shimojo (2002) observaron que el número de sonidos presentados modula el número de flashes percibidos sólo si la asincronía entre ambos no supera los 100 milisegundos (ms). Sambo y Foster (2008) reportaron un incremento en la amplitud de los potenciales evocados asociados a áreas somatosensoriales cuando un estímulo visual era presentado cercano a la localización de un estímulo táctil. Por su parte, Senkowski, Talsma, Grigutsch, Herrmann y Woldroff (2007) encontraron que la interacción entre estímulos visuales y auditivos, medida tanto conductualmente como a través de frecuencias gamma, se hacía más robusta conforme el nivel de asincronía entre la presentación de los estímulos se hacía menor. Macaluso, George, Dolan, Spence y Driver (2004) disociaron los aspectos temporales y espaciales en el procesamiento del habla, mostrando como áreas ventro-occipitales se asociaban con la presentación de estímulos audiovisuales

sincronizados, mientras que áreas dorso-occipitales respondían ante estímulos bisensoriales originados desde la misma localización externa.

1.3. Factores determinantes: cognitivos

Como se ha puesto de manifiesto en el apartado anterior, el estudio de las características de los estímulos, los sistemas sensoriales y las bases neurales que subyacen al procesamiento multisensorial han sido ampliamente tratados a lo largo de los años. Mientras, la influencia que otros procesos cognitivos ejercen a la hora de combinar información proveniente de diferentes modalidades sensoriales no ha empezado a estudiarse hasta finales del siglo pasado, con un gran auge en la última década de este siglo (véase Driver y Spence, 1998; Spence y Driver, 2004 para una revisión sobre el papel de la atención).

La atención es sin duda el proceso cognitivo más estudiado en relación al procesamiento multisensorial, y en el que nos vamos a centrar en apartados posteriores debido a su relevancia para la tesis que nos ocupa. Sin embargo, no ha sido el único factor importante investigado. Así, por ejemplo, se ha demostrado como la discriminación de estímulos visuales mejora tras haber sido emparejados con estímulos auditivos irrelevantes (véase Murray, Michel, Grave de Peralta, Ortigue, Brunet, Andino y Schnider, 2004 para asociaciones audiovisuales realistas - v.gr., el dibujo de una campana acompañado del sonido “dong” -; también Zangenehpour y Zatorre, 2010 para la asociación de estímulos audiovisual arbitrarios), e incluso como las áreas neurales implicadas en la percepción multisensorial de dichos estímulos se activan ante la presentación del estímulo visual de manera aislada. De este modo vemos como la percepción se ve modula por procesos cognitivos tales como el aprendizaje y la memoria relacionados con estímulos multisensoriales.

1.4. Paradigmas para el estudio del procesamiento multisensorial: Conflicto

Varios son los paradigmas que se han utilizado para el estudio del procesamiento multisensorial. Entre otros tenemos: 1) el *paradigma de señalización* (Posner, 1980) que ha permitido demostrar como una señal en una modalidad sensorial mejora la percepción de estímulos en otra modalidad sensorial (v.gr., Schmitt, Postma y Haan, 2000, 2001); 2) el *paradigma de doble tarea* que ha permitido observar como la capacidad de los recursos cognitivos puede influir en el procesamiento de estímulos multisensoriales (v.gr., Diver y Spence, 1994); 3) el *paradigma de redundancia sensorial* que muestra como la presentación de estímulos en más de una modalidad sensorial mejora la percepción de los mismos en relación a su presentación de forma aislada (v.gr., Stein, London, Wilkinson y Price, 1996; Vroomen y de Gelder, 2000); 4) el *paradigma del juicio del orden temporal* que permite discernir el procesamiento temporal entre las distintas modalidades sensoriales, ayudando a entender cuáles son las características sincrónicas necesarias para que se produzca una experiencia multisensorial (v.gr., Spence, Shore y Klein, 2001; Vroomen y Keetels, 2010); 5) el *paradigma de correspondencia* que estudia las que características que de modo “innato” tienen una correspondencia entre las distintas modalidades sensoriales, como los sonidos agudos con objetos brillantes y/o situados en posiciones altas y los sonidos graves con objetos sombríos y/o situados en posiciones bajas (v.gr., Marks, 2004; Spence, 2011).

Uno de los paradigmas ampliamente utilizado y que es central para esta tesis es el *paradigma de conflicto*. Mediante este paradigma se crean situaciones artificiales de conflicto perceptivo con el objetivo de conocer como el sistema cognitivo resuelve la incongruencia entre las características de los estímulos para así conformarse una percepción ajustada de la realidad del objeto. Son muchos los modos en que se puede crear el conflicto multisensorial en el laboratorio. Por ejemplo, Warren, Welch y McCarthy (1981), o más recientemente Alais y Burr (2004), mostraron como ante la presentación simultánea de información visual y auditiva en distintas localizaciones espaciales se

producía un efecto de captura auditiva por parte de la información visual - i.e., *efecto del ventrílocuo* -, de tal manera que se percibía el sonido como proveniente de la localización en que se presentaban los estímulos visuales (véase Caclin, Soto-Faraco, Kingstone y Spence, 2002 para el efecto entre las modalidades táctil y auditiva). Además de con estímulos estáticos, como acabamos de ver, este efecto también se ha estudiado con estímulos en movimiento denominándose *captura dinámica*. Por ejemplo, en un estudio llevado a cabo por Sanabria, Soto-Faraco y Spence (2007a) dos luces situadas a relativa distancia se encendían secuencialmente dando lugar a la percepción de un estímulo visual que se desplazaba (i.e., movimiento aparente). Del mismo modo, la percepción de movimiento de un sonido se creaba mediante altavoces situados en distintas localizaciones. Cuando la dirección de movimiento de estímulos visuales y auditivos difería, los participantes informaban que la dirección del movimiento auditivo era la misma que la de la cadena de estímulos visuales. Otro modo de crear conflicto pasa por presentar distinto número de estímulos en modalidades sensoriales diferentes, como hicieron Shams, Kamitani y Shimojo (2000) en *la ilusión del doble flash*, según la cual el número de flashes percibidos depende del número de sonidos realmente presentados.

No sólo características físicas simples como las enumeradas hasta ahora han sido manipuladas para crear conflicto, sino también características más complejas como la información fonética o semántica han permitido estudiar el modo en que el cerebro integra la información multisensorial. Un buen ejemplo es el ya mencionado *efecto McGurk* (McGurk y McDonald, 1976) en el que la incongruencia entre la información fonética visual y auditiva da lugar a la percepción de un nuevo fonema. Beauchamp, Lee, Argall y Martin (2004) por su parte, estudiaron como información audiovisual semánticamente congruente (v.gr., la imagen de un teléfono acompañada del sonido “ring”) provocaba una mayor activación en áreas responsables del procesamiento multisensorial (i.e., Surco temporal superior o Giro temporal medio) que estimulación semánticamente incongruente (v.gr., la imagen de un teléfono acompañada del sonido “bang”).

Dado que en el mundo lo habitual es que la información que proviene de un mismo evento, aunque llegue a través de distintas modalidades sensoriales, sea congruente entre dichas modalidades en propiedades tales como posición, duración, sincronía..., la relevancia del paradigma de conflicto viene dada porque la variación de esta congruencia en alguna propiedad de los estímulos permite discernir la naturaleza de la combinación de la estimulación, mostrando qué modalidad sensorial tiene más peso a la hora de guiar la integración, cuáles son las características importantes para que se dé la integración, etc. Mientras la relevancia de otros paradigmas no se centraba en la naturaleza de la combinación, sino en la ventaja de que el sistema cognitivo pueda integrar la información de distintos sentidos para obtener una percepción y un procesamiento mejor del que se tendría basado sólo en la estimulación unisensorial. Ejemplos de estos estudios son los resultados que muestra como somos más rápidos en reaccionar cuando la información nos llega por dos modalidades sensoriales (v.gr., reaccionar hasta 600 ms más rápido si ante la inminente colisión con un coche se recibe un aviso audiotáctil, frente a un aviso unisensorial, v.gr., Ho, Reed y Spence 2007; Ho y Spence, 2008); o como se produce una mejora en la respuesta a una modalidad, cuya señal no es lo suficientemente buena, cuando se acompaña de información en otra modalidad (v.gr., cuando hay mucho ruido de fondo y no se puede percibir claramente el sonido de alguien que te habla, la posibilidad de poder ver el movimiento de los labios de la persona mejora mucho la percepción del habla; v.gr., Ross, Saint-Amour, Leavitt, Javitt y Foxe, 2007).

1.5. Otras aproximaciones: Nivel de procesamiento

Un enfoque con gran peso en el estudio del procesamiento multisensorial, y relevante para la tesis que nos ocupa, es el tópico de a qué nivel se está dando la relación entre información proveniente de distintas modalidades sensoriales. Esta relación puede ocurrir a nivel perceptivo, es decir, un procesamiento sensorial previo a la selección de respuesta y ejecución de la misma, o a nivel post-perceptivo relacionado con las etapas

de decisión y ejecución de respuesta (véase Aschersleben, Bachmann y Musseler, 1999; Bertelson y de Gelder, 2004). La importancia de advertir a qué nivel se está dando el procesamiento multisensorial radica en conocer si la información que recibe el sistema se combina e interactúa a niveles tempranos, generando un evento multisensorial que continuará hacia los siguientes estadios de procesamiento (v.gr., selección de respuesta), o si por el contrario no hay modulaciones a nivel temprano debidas a la presencia de información multisensorial - frente al procesamiento de información unisensorial -, y lo que ocurre es que procesos más tardíos serán sesgados por la presencia de dicha información multisensorial.

Varios son los modos para poder profundizar en la temática de los niveles de procesamiento. A nivel conductual, la Teoría de Detección de Señales (TDS; véase Green y Swets, 1966; Macmillan y Creelman, 1991) permite el cálculo estadístico de un índice de sensibilidad o discriminabilidad (d') de los estímulos, que se asocia con el procesamiento perceptivo, así como un índice de sesgo en la respuesta o criterio (β o c), que indica cuan laxo o cuan estricto es el criterio que la persona tiene a la hora de dar la respuesta, y que por tanto se asocia con el procesamiento post-perceptivo. Sanabria, Spence y Soto-Faraco (2007b), presentando información visual, auditiva y audiovisual congruente e incongruente, encontraron que la sensibilidad sensorial (d') disminuía ante la presencia de estimulación en una segunda modalidad independientemente de que ésta fuese congruente o incongruente. Por el contrario, el criterio de respuesta (β o c) se hacía menos laxo ante estimulación audiovisual congruente y más estricto ante estimulación audiovisual incongruente. Estudios como éste ponen de manifiesto como la interacción multisensorial puede darse a ambos niveles de procesamiento de manera independiente uno del otro.

Otro modo de observar a qué nivel tiene lugar el procesamiento multisensorial es a través de técnicas de electroencefalografía y magnetoencefalografía. Estas técnicas típicamente se caracterizan por su alta resolución temporal, de modo que el momento

temporal en que se detecta la modulación puede ser tomado en consideración para indicar que estamos ante un procesamiento de tipo perceptivo (i.e., hasta los 200 ms aproximadamente; véase Rugg y Coles, 1996; Luck, 2005) o ante un procesamiento post-perceptivo (i.e., a partir de los 200 ms). Como ejemplo tenemos el estudio de Mishra, Martinez, Sejnowski y Hillyard (2007) en el que analizando la actividad eléctrica cerebral asociada con la ilusión del doble flash (i.e., percepción equivocada de dos flashes cuando un único flash se presenta acompañado de dos sonidos), observaron que entre los 100 y 130 ms los potenciales asociados al procesamiento visual y auditivo mostraban una mayor amplitud cuando los participantes habían percibido la ilusión frente a cuando no la habían percibido. Por su parte, Raij, Uutela y Hari (2000) llevaron a cabo un estudio de magnetoencefalografía y observaron que la presentación de grafemas acompañados de sus fonemas producía mayor actividad entre los 380 y 540 ms que la presentación de grafemas acompañados de un fonema diferente.

Los estudios de neuroimagen (v.gr., Imagen por Resonancia Magnética Funcional, IRMf) también nos dan información acerca de los niveles de procesamiento, dado que permiten observar si las áreas implicadas en el procesamiento de la información se corresponden con áreas de bajo o de alto nivel. De manera que la actividad en áreas sensoriales primarias y de asociación indicaría procesamiento a nivel perceptivo. Ejemplo de ello es el estudio que realizaron Watkins, Shams, Josephs y Rees (2007) en el que encontraron que la actividad de la corteza visual primaria (V1) aumenta cuando se percibe la ilusión del doble flash y disminuye cuando la presentación de dos flashes con un solo sonido produce la ilusión contraria, es decir la percepción errónea de un solo flash.

En suma, este apartado pone de manifiesto que el procesamiento multisensorial puede tener lugar en distintos estadios del procesamiento de la información.

1.6. Mecanismos neurales

Muchos esfuerzos se han puesto en estudiar y conocer el entramado neural en que se respalda el procesamiento multisensorial y todos los efectos que éste conlleva. Muchos avances se han hecho desde que Stein y Meredith (1993) comenzaron estudiando a nivel de registro unicelular aquellas neuronas que respondían diferencialmente a estimulación bimodal, e incluso trimodal, en las células de los colículos superiores del gato hasta la actualidad. A continuación se pretende sintetizar del modo más claro y conciso los hallazgos más relevantes sobre el estudio de la circuitería neural del procesamiento multisensorial.

Desde un punto de vista clásico se han definido dos tipos de estructuras subyacentes al procesamiento multisensorial: estructuras subcorticales y áreas multisensoriales de alto nivel. Sin embargo, recientemente las áreas corticales de bajo nivel (i.e., áreas sensoriales primarias y secundarias), consideradas hasta el momento áreas de procesamiento específicas de modalidad, se han añadido como tercer tipo de estructura tras las evidencias obtenidas de estudios anatómicos y funcionales.

A nivel subcortical las áreas implicadas en el procesamiento multisensorial son, fundamentalmente, las siguientes:

Los *colículos superiores*, que integran información de distintas modalidades sensoriales, como ya habían reflejado Stein y Meredith (1993) en los efectos neurales de superaditividad. Esta integración se respalda en la gran cantidad de inputs ascendentes y descendentes que reciben de áreas unisensoriales (v.gr., May, 2006; Wallace, Meredith y Stein, 1993). Como muestra de ello está un estudio llevado a cabo por Alvarado, Stanford, Vaughan y Stein (2007) en el que se demostró que cuando la conexión entre áreas visuales y colículos superiores estaba ausente se mantenía intacto el procesamiento de características unisensoriales, pero se veía imposibilitada la integración de estímulos multisensoriales.

El *tálamo*, que constituye el punto de relevo para que la información sensorial de las principales modalidades llegue al córtex a través de las proyecciones tálamo-corticales. Por ejemplo, la información visual llega de las células de la retina al córtex visual primario a través del núcleo geniculado lateral del tálamo, mientras la información auditiva tiene el relevo en el núcleo geniculado medial del tálamo antes de llegar al córtex auditivo primario. Pero el tálamo no sólo es una estación de relevo unisensorial, sino que hay ciertos núcleos que se encargan de procesar información multisensorial antes de enviarla a las cortezas primarias, así como de recibir inputs multisensoriales que ya han sido integrados en otras áreas. Estos núcleos se encuentran situados en la parte posterior del tálamo, siendo los más estudiados el suprageniculado, el limitante y el pulvinar, demostrándose la existencia de neuronas que responden a actividad visual, auditiva y somatosensorial (véase Mchaffie, Fuentes-Santamaría, Alvarado, Fuentes-Farías, Gutiérrez-Ospina y Stein, 2002 para una revisión)

Otra estructura a tener en cuenta es el *claustró*, cuya actividad está bastante relacionada con la corteza insular. El claustró cumple un papel primordial en el emparejamiento de características similares entre información de distintas modalidades sensoriales (v.gr., Hadjikhani y Roland, 1988; Banati, Goerres, Tjoa, Aggleton y Grasby, 2000)

Las áreas de la corteza cerebral que tienen implicación en el procesamiento multisensorial, agrupadas en áreas de alto nivel y bajo nivel, son las siguientes:

El *surco temporal superior* gracias a las conexiones bidireccionales que tiene con las cortezas auditiva, visual y somatosensorial primarias (v.gr., Cusick, 1997; Padberg y cols., 2003; Schmahmann y Pandya, 1991), así como por la presencia de neuronas multisensoriales (v.gr., Barraclough y cols., 2005) responde a información proveniente de estas tres modalidades sensoriales. El efecto típicamente observado es un aumento de la actividad en este área si se presentan, por ejemplo, estímulos audiovisuales frente a la

presentación de estimulación de cada modalidad aislada como ocurre en el estudio realizado por Beauchamp, Lee, Argall, y Martin (2004).

Además hay que destacar que el surco temporal superior juega un papel primordial en la integración de información audiovisual, fundamentalmente lingüística, mostrando, por ejemplo, una mayor activación cuando palabras presentadas auditivamente son congruentes con movimientos labiales o con la presentación visual de letras (v.gr., Calvert y cols., 2000 o Raji, Uutela y Hari, 2000, respectivamente)

El *surco intraparietal*, que integra información visual, vestibular y táctil (v.gr., Bremmer, Schlack, Shah, Zafiris, Kubischik, Hoffman y cols., 2001) además de recibir proyecciones de áreas multisensoriales (v.gr., Lewis y Van Essen, 2000).

La *ínsula*, principalmente su parte posterior, que está fuertemente interconectada con los núcleos multisensoriales del tálamo (v.gr., Hackett, De La Mothe, Ukbirt, Karmos, Smiley y Schroeder, 2007), y que contiene neuronas que responden a estimulación visual, auditiva y somatosensorial (v.gr., Bieser, 1998).

Por su parte, la implicación de la *corteza frontal* (i.e., córtex prefrontal y córtex premotor) en el procesamiento multisensorial todavía no es muy bien conocida. A grandes rasgos parece asociarse con características puntuales de la tarea como, por ejemplo, el procesamiento de nuevas asociaciones multisensoriales (v.gr., Giard y Peronet, 1999) o la ejecución de tareas de emparejamiento multisensorial (v.gr., Bushara y cols., 2001; Banati y cols., 2000). Aunque también se ha demostrado que, principalmente, el córtex premotor juega un papel importante en el procesamiento de movimiento entre estímulos multisensoriales (v.gr., Bremmen, Schlack, Shah, Zafiris, Kubischik, Hoffmann, Zilles y Fink, 2001) gracias a las neuronas que responden a estímulos visuales, auditivos y táctiles y a las conexiones con áreas parietales (v.gr., Luppino, Murata, Govoni, y Matelli, 1999). También tiene una importante participación en la configuración del espacio peripersonal con el objetivo de interactuar con él, gracias a la sensibilidad que muestran las neuronas multisensoriales a las distintas localizaciones de los estímulos (v.gr.,

Graziano, Reiss y Gross, 1999). De hecho, estudios como el de Ládavas, Pellegrino, Franè y Zeloni (1999) muestran como lesiones en áreas premotoras modulan la integración de información viso-táctil en el espacio peripersonal, pero no en el espacio extrapersonal.

En cuanto a las cortezas sensoriales los estudios demuestran que gracias a las proyecciones directas que hay entre las áreas primarias (v.gr., Rockland y Ojima, 2003; Cappe y Barone, 2005) se producen modulaciones en los córtex asociados a una modalidad concreta cuando se presenta información en otras modalidades. Por ejemplo, Kayser, Petkov y Logothetis (2009) encontraron evidencias de influencia de la información visual y somatosensorial en la corteza auditiva primaria, y Calvert y colaboradores (1999) mostraron como tanto el córtex auditivo como el visual muestran mayor activación cuando se ve y escucha al mismo tiempo a una persona hablando, que la actividad de cada una de esas áreas cuando sólo se ve o sólo se escucha a la persona. Por supuesto la densidad de la conexiones entre áreas sensoriales primarias es mucho menor que la cantidad de conexiones que hay entre las áreas de alto nivel, pero no por ello deben dejar de tenerse en cuenta.

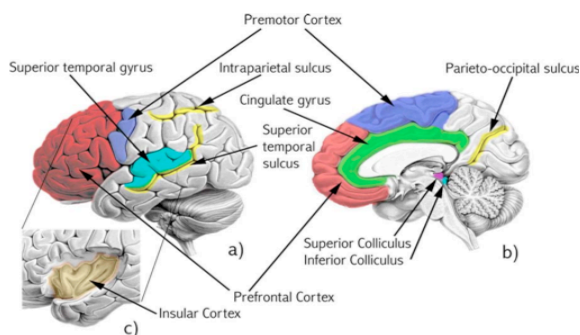


Figura1. Áreas implicadas en el procesamiento multisensorial en el cerebro humano.
(Tomado de Calvert y Thesen, 2004).

Se pone de manifiesto en este apartado como la red neural encargada del procesamiento multisensorial es muy extensa y compleja, abarcando estructuras corticales y subcorticales así como todas las posibles conexiones entre ellas (v.gr., 36

Ghazanfar y Schroeder, 2006). El funcionamiento de esta red se basa principalmente en dos tipos de actividad neural, bien la presencia de neuronas multisensoriales que se encargan de integrar información de distintas modalidades sensoriales en cada una de las áreas de procesamiento, o bien se trata de una modulación de unas áreas sobre otras gracias a las proyecciones que transmiten información unisensorial o multisensorial.

Por último, comentar que los distintos niveles anatómicos implicados en el procesamiento multisensorial parecen estar relacionados con una finalidad funcional, pudiéndose relacionar cada nivel con un tratamiento de información diferente. Así, a nivel subcortical (v.gr., tálamo y colículos superiores) la naturaleza de la información tiene que ver con características estructurales como puede ser la alineación espacial. Mientras que a nivel de áreas corticales se asocia con procesos de identidad del estímulo. A nivel de áreas primarias se identificarán características como tamaño, orientación, tono, etc, al tiempo que en áreas de nivel superior entra en juego la identificación de características semánticas así como la selección de respuesta.

Resumiendo, debe ser tenido en cuenta que el procesamiento multisensorial es el producto de un red global que gracias a la especificación funcional y anatómica permite una experiencia perceptiva coherente del mundo que nos rodea.

2. Atención y procesamiento multisensorial

Nuestro sistema cognitivo está constantemente expuesto a una gran cantidad de estimulación sensorial. Como hemos visto, la combinación de dicha información es crucial a la hora de componer una percepción verídica del mundo para así interactuar de un modo adecuado. Retomando el ejemplo que presentábamos al comienzo, la pregunta que de manera automática se plantea es ¿cómo el cerebro identifica y selecciona aquella información que debe combinarse de entre toda la información presente? Es decir, ¿cómo

la información visual proveniente de los movimientos de los labios de los jóvenes se integra con la información auditiva proveniente del sonido de las palabras, y no con el sonido del trino de los pájaros? Hemos visto que características como la localización espacial y la sincronía temporal pueden ayudar a elaborar el percepto final, así en este caso el hecho de que el lugar espacial del que llega el sonido coincida con la posición de los labios de los jóvenes, o que el momento en que se comienzan a mover los labios y a producirse el sonido sea el mismo apoya la idea de que ambas estimulaciones pertenezcan al mismo evento (i.e., asunción de unidad; v.gr., Welch, 1999; Vatakis y Spence, 2007). Pero ¿qué ocurre con el resto de estimulación que está llegando a nuestros sentidos y que también puede coincidir espacial y temporalmente con el movimiento labial o con el sonido de las palabras? ¿Cuál es el mecanismo o proceso encargado de resolver esta situación de sobreestimulación?

La atención se convierte en el candidato perfecto para ocuparse de seleccionar la información del ambiente que es relevante para una correcta interacción con el mundo que nos rodea. Y por tanto el foco de interés se centrará en ver como los mecanismos atencionales se combinan con las propiedades de los estímulos para dar lugar a una percepción coherente de la experiencia multisensorial.

2.1. Atención

Entendemos la atención como la función cognitiva encargada de, por un lado, seleccionar de toda la información que se recibe del ambiente aquellas características que son relevantes para el objetivo actual, y por otro lado, de ensamblar dichas características relevantes en conjuntos estímulares coherentes.

Hemos visto en apartados anteriores como determinadas características estructurales de la información (v.gr., (a)sincronía o localización) por sí mismas ayudan al ensamblaje correcto del percepto final - i.e aquella información que se presenta cercana en tiempo y espacio tenderá a unirse, mientras aquella lejana en espacio y tiempo tenderá

a considerarse perteneciente a distintos objetos -. Sin embargo, esta facilitación no puede dar cuenta de la selección correcta de la información en todas las situaciones, puesto que, como hemos visto, en el ambiente información irrelevante (v.gr., trino de los pájaros en nuestro ejemplo) puede estar presente de manera sincrónica y en cercana posición espacial con información relevante (v.gr., movimiento de los labios), situación ésta que se ve maximizada en los paradigmas de conflicto, centrales para nuestra investigación. Aquí es donde los procesos atencionales guiados por objetivos concretos desempeñan su papel a lo largo de las distintas etapas del procesamiento de la información.

En un primer momento toda la información que está llegando a nuestros sentidos es procesada a nivel perceptivo, si dichas características son coherentes entre ellas y no existe más información que compita por el procesamiento la atención es atraída (v.gr., atención guiada por los estímulos, procesamiento de abajo-arriba) y se ocupa de crear una representación coherente mediante la unión de todas las características presentes, como pueden ser forma, tamaño, color, localización, etc.

Sin embargo, como se mencionaba en párrafos anteriores, lo más común en el medio ambiente y fundamentalmente en el laboratorio (v.gr., paradigma de conflicto) es que se dé un cierto nivel de competición entre las representaciones de los estímulos que el sistema debe procesar. En este caso puede ocurrir que la resolución de dicha competición esté guiada por las propias características de los estímulos (procesamiento de abajo-arriba) que, por ejemplo, harían unos estímulos más salientes que otros o desencadenarían esquemas de procesamiento fuertemente asociados a dichas características. O bien mediante mecanismos guiados por la meta (o procesamiento de arriba-abajo) que se encargarían de sesgar el procesamiento de la información hacia aquellas características y/o representaciones que son relevantes para el control del comportamiento (i.e., *Teoría de la competición sesgada o integrada*; Desimone y Duncan, 1995). Por ejemplo, Duncan, Martens y Ward (1997) encontraron que la ejecución de una tarea de identificación tanto visual como auditiva se ve mejorada cuando las instrucciones

piden atender sólo a una de dos cadenas de estímulos frente a cuando se trata de una tarea de atención dividida.

Además de seleccionar de toda la información presente aquella que es relevante, la atención también se ocupa de combinar correctamente dicha información. Esta idea de que la atención se requiere para la integración de rasgos en una unidad (v.gr., Treisman y Gelade, 1980; Treisman, 1998) se pone de manifiesto en estudios como el llevado a cabo por Treisman y Schmitd (1982) en el que demuestran como cuando la atención está distribuída por la totalidad de un display compuesto por diferentes letras y números de distintos colores situadas en varias localizaciones, las personas tienden a unir de manera equívoca - *conjunciones ilusorias* - los rasgos letra-color-localización, frente a cuando se les permite focalizar la atención, por ejemplo, en una localización concreta. Igualmente, se obtiene una mejora en la ejecución para estímulos caracterizados por varios rasgos (i.e., cruz formada por línea azul y otra verde) frente a estímulos caracterizados por sólo un rasgo (i.e., línea verde), cuando se tiene información sobre la localización probable de aparición del estímulo, es decir cuando se puede focalizar la atención en una localización frente a cuando la atención está distribuida por a totalidad del display (véase Treisman, 2006).

Estudios como el llevado a cabo por Crump, Vaquero y Milliken en 2008, muestran la importancia de la atención tanto en su función de seleccionar información relevante, como en la de crear conjunciones adecuadas. Estos autores asociaron la forma del estímulo con la probabilidad de que éste fuese congruente en una tarea tipo Stroop (Stroop, 1935). Encontraron que sólo cuando la característica forma debía ser selecciona por ser relevante, ya que se debía llevar el recuento de las distintas formas presentadas, se producía la unión entre forma y proporción de congruencia como pertenecientes al mismo objeto.

La atención, por lo tanto, se encarga de seleccionar la información relevante, ya sea a través de un procesamiento guiado por los estímulos o de un procesamiento guiado por

40

las metas, y de integrar dicha información en una percepción final que nos permite interactuar con el mundo que nos rodea adaptando nuestro comportamiento de acuerdo a los objetivos del momento.

2.2. Primeras aproximaciones

En un primer momento el procesamiento multisensorial se concebía como un proceso automático guiado por las características físicas de los estímulos, así como por las características estructurales y funcionales de los sentidos. Siendo así no había lugar a que procesos cognitivos como la atención pudiesen influir en el proceso y resultado de la experiencia multisensorial. Esta hipótesis pre-atentiva del procesamiento multisensorial se apoyaba en estudios como el de Driver (1996, Exp. 2) en el que pedía a los participantes repetir la información auditiva presentada por un altavoz mientras se ignoraba la información auditiva presentada simultáneamente en otro altavoz separado espacialmente. El porcentaje de respuestas correctas disminuía cuando se presentaban al mismo tiempo movimientos labiales congruentes con la información auditiva distractora. De manera que se defendía que la integración audiovisual (v.gr., efecto del ventrílocuo) se daba de manera automática y previa a que los procesos de selección auditivos pudiesen tener lugar (véase también Bertelson, Vroomen, de Gelder y Driver, 2000; Vroomen, Bertelson y de Gelder, 2001 para procesos de atención visual).

Al mismo tiempo otra serie de estudios retaban esta postura estudiando principalmente como cambios atencionales en una modalidad sensorial pueden afectar a la orientación y/o ejecución en otras modalidades (véase Spence y Driver, 2004 para una revisión). Así, un estudio llevado a cabo por Spence y Driver (1996) mostró como la orientación de la atención a una localización concreta no sólo mejoraba la discriminación del estímulo visual objetivo, sino también, aunque en menor medida, la discriminación de un estímulo auditivo irrelevante para la tarea. Schröger y Eimer (1997) encontraron el mismo patrón de resultados en una tarea go/no-go en la que se medía la actividad

electroencefalográfica (EEG) - i.e., mayor amplitud en las ondas asociadas al estímulo visual objetivo, así como un incremento de menor magnitud en la respuesta a un estímulo auditivo irrelevante cuando se presentaban en la localización atendida -.

Estos últimos estudios presentados ponen de manifiesto la existencia de un vínculo entre la atención a las distintas modalidades sensoriales, pero no descartan la posibilidad de que la interacción multisensorial ocurra pre-atentivamente, ya que el objeto real de estudio era como la atención (generalmente guiada por una señal) en una modalidad sensorial modula la percepción de un estímulo subsecuente en otra modalidad sensorial, lo que se conoce como *atención crossmodal* (véase Koelewijn, Bronkhorst y Theeuwes, 2010 para una revisión actualizada de este tipo de estudios) .

2.3. Perspectiva actual

Un cambio en la aproximación empleada hasta el momento era necesario para poder comparar los resultados con los estudios que demostraban que el procesamiento multisensorial era previo e inmune a la atención. Para ello se comenzó a utilizar la misma estrategia que en esas investigaciones, esto es, partiendo de efectos multisensoriales conocidos (o nuevos fenómenos multisensoriales corroborados) aplicar manipulaciones atencionales (v.gr., atender a una modalidad sensorial o a una localización espacial) y comprobar si dichos efectos multisensoriales varían, desaparecen o se incrementan. Profundizando de este modo en el estudio del papel de la atención en el procesamiento multisensorial *per se*.

Dentro de esta aproximación encontramos estudios como el de Talsma y Woldorff (2005) donde utilizando una tarea oddball en la que se presentaban a ambos lados de una pantalla estímulos visuales, auditivos y audiovisuales, pidieron a los participantes que detectasen los cambios en la duración de dichos estímulos sólo en uno de los lados de la pantalla. Para asegurarse que los estímulos utilizados daban lugar a un fenómeno multisensorial utilizaron el método propuesto a nivel unicelular por Meredith y Stein (1983)

y adaptado por primera vez a un estudio sobre encefalografía humana por Giard y Peronnet (1999), en el que se compara la actividad ante la presencia de estímulos multisensoriales con la suma de la actividad ante la presencia separada de estímulos visuales y auditivos, indicando un resultado multisensorial cuando hay una diferencia entre ambas actividades [$AV \neq A+V$]. Observaron que los potenciales evocados asociados a los estímulos audiovisuales atendidos eran de mayor amplitud que los asociados a los no atendidos en momentos temporales diferentes a lo largo el flujo de procesamiento. Por lo que concluyeron que el procesamiento multisensorial es sensible a procesos atencionales en distintas etapas del procesamiento. Resultados similares obtuvieron Mishra, Martínez y Hillyard (2010) quienes partiendo de la ilusión del doble flash (Shams y cols., 2000) aplicaron también una manipulación de la atención espacial. Además, en este estudio se observaba también que el efecto atencional sobre el procesamiento multisensorial ocurría en áreas de modalidad específica (i.e., córtex estriado visual) y en áreas heteromodales (i.e., córtex temporal superior).

En esta línea de investigación Busse, Roberts, Crist, Weissman, y Woldorff (2005) llevaron a cabo un estudio en el que pretendían comprobar si el efecto del ventrículo era inmune a la atención, como se había defendido anteriormente (véase Driver, 1996). En este estudio se presentaban estímulos visuales a ambos lados de la pantalla y un estímulo auditivo central, y se pedía a los participantes que atendiesen nuevamente sólo a los estímulos visuales presentados en uno de los lados de la pantalla. Los resultados mostraron, a nivel de actividad cerebral (i.e., EEG, IRMf), un incremento del efecto del ventrículo cuando el sonido se presentaba acompañando a un estímulo visual atendido frente a un estímulo visual no atendido. Concluyeron que el procesamiento multisensorial sí puede estar modulado por procesos atencionales. Ciertas diferencias en el diseño de ambos experimentos (i.e., Driver, 1996 y Busse y cols., 2005) pueden explicar las distintas conclusiones que obtienen. Driver (1996) presentaba dos cadenas auditivas simultáneamente y en el centro de la pantalla estímulos visuales neutros o congruentes

con la información auditiva irrelevante. El resultado esperado para defender la hipótesis atencional era que como los participantes estaban atendiendo al mensaje auditivo objetivo para indicar su localización espacial, la presentación de información visual congruente con el mensaje irrelevante no tendría ninguna influencia en el resultado. Sin embargo, la presentación de la información visual irrelevante disminuía el número de aciertos a la hora de localizar el mensaje auditivo objetivo. Driver (1996) defendía que esto ocurría porque la información visual y auditiva irrelevante se integraban de manera automática independientemente de las instrucciones dadas. Mientras en el estudio de Busse y colaboradores (2005) la misma información visual era presentada a ambos lados de la pantalla y solamente las instrucciones atencionales guiaban la integración de la información central auditiva que se presenta en ocasiones acompañando a los estímulos atendidos y en ocasiones a los no atendidos. Dos modos diferentes de relación entre atención y procesamiento multisensorial se están viendo en ambos estudios. En el caso de Driver (1996) la redundancia de la información multisensorial, aunque irrelevante para la tarea, hace a los estímulos lo suficientemente salientes como para que de por sí mismos capten la atención en un procesamiento guiado por los estímulos (i.e., de abajo-arriba) dando lugar a un empeoramiento en la localización del estímulo auditivo objetivo. Por su parte, en el estudio de Busse y colaboradores (2005) se puede observar como el procesamiento de los estímulos visuales y auditivos ocurría tanto en la condición de atendido como de no atendido, puesto que como comentábamos la información redundante atrae la atención, pero además las instrucciones atencionales dan lugar a un incremento de este efecto, puesto que la atención guiada por el objetivo de la tarea (i.e., de arriba-abajo) mejora el procesamiento pudiendo así modular el procesamiento multisensorial.

Otros estudios defienden este enfoque bidireccional entre la atención y el procesamiento multisensorial (véase Talsma, Senkowski, Soto-Faraco y Woldorff, 2010). Estudios como los Vroomen y de Gelder (2000) o Van der Burg, Olivers, Bronkhorst y

Theeuwes (2008) son ejemplos de un procesamiento atencional de abajo-arriba puesto que la presentación simultánea de estimulación auditiva mejora la detección, en el primer estudio, y la búsqueda, en el segundo, de los estímulos visuales, es decir la redundancia sensorial, provoca el cambio en la atención que mejora el procesamiento de los estímulos. Mientras la relación entre la atención y el procesamiento multisensorial desde una perspectiva de arriba-abajo se pone de manifiesto en estudios de atención espacial como el llevado a cabo por Sanabria, Soto-Faraco y Spence (2007a) en el que el efecto de captura dinámica audiovisual se veía disminuido cuando la estimulación ocurría en una localización atendida; o el estudio de Farihall y Macaluso (2009) en el que se mostró activación de áreas asociadas al procesamiento audiovisual (i.e., surco temporal superior y colículos superiores) sólo cuando se atendía a estimulación audiovisual congruente frente a cuando no era atendida; así como el estudio de Mishra, Martínez y Hillyard (2010) donde los potenciales evocados asociados a la ilusión del doble flash veían incrementada su amplitud en la localización atendida frente a la no atendida.

Talsma y colaboradores (2010) conjugan el amplio abanico de resultados en torno a un marco polifacético de la relación entre atención y procesamiento multisensorial. Así, en línea con el concepto de atención que desarrollamos en el apartado 2.1, defienden la posibilidad de, por una parte, un procesamiento guiado por los estímulos según el cual características estructurales como, por ejemplo, la saliencia o el alineamiento espaciotemporal, llevarán a la integración de los estímulos lo que en segunda instancia atraerá la atención, y por otra parte, un procesamiento guiado por las metas que entrará en juego cuando las representaciones de los estímulos entren en competición y se requiera de un objetivo que priorice el procesamiento de la información relevante.

La idea planteada en este marco teórico bidireccional encuentra apoyo a nivel anatómico en otras aproximaciones que describen las distintas formas de interacción que se pueden dar entre atención y percepción. De este modo, la atención guiada por los estímulos se correspondería con lo que Corbetta, Patel y Shulman (2008) denominaron

red neural ventral temporoparietal frontal, también conocida como red ventral de orientación en el modelo atencional de Posner (v.gr., Petersen y Posner, 2012). Esta red, que sería la responsable de detectar los estímulos sensorialmente relevantes (v.gr., salientes) y orientar la atención hacia ellos, implicaría la actividad de áreas tales como la unión temporoparietal y el córtex frontal ventral. Mientras la atención guiada por las metas se correspondería con la red dorsal frontoparietal - Corbeta y cols., 2008 - y la red ventral de orientación - Petersen y Posner, 2012 -. Esta segunda red, compuesta por áreas como los surcos precentral e intraparietal, seleccionaría los inputs sensoriales relevantes a partir de la información preexistente y metas internas, y asociaría los estímulos relevantes con las respuestas (v.gr., Rushworth y cols., 2001). Las zonas anatómicas descritas son consistentes con las áreas frontoparietales que conforman el sistema de múltiple demanda, que según Duncan, se encarga de sesgar la actividad hacia la información crítica o relevante (v.gr., Duncan, 2013).

2.4. Niveles de procesamiento

En el apartado anterior hemos visto como la evidencia empírica ha mostrado como los procesos atencionales pueden modular el procesamiento de estímulos multisensoriales generando un percepto final diferente en función de la distribución de la atención. Sin embargo, esto no es suficiente para comprender completamente la relación entre atención y procesamiento multisensorial. El nivel de procesamiento en el cual ambos procesos interactúan es un aspecto importante, y que ha sido tenido en cuenta en estudios que se centran en estudiar si dicha interacción ocurre a nivel perceptivo o post-perceptivo, como ocurre con el procesamiento multisensorial *per se*. Esta distinción implica conocer si realmente la atención puede influir en cómo los estímulos se combinan para generar el percepto final, o si por el contrario el procesamiento perceptivo no se modula y el papel de la atención consiste en sesgar la decisión a la hora de seleccionar y dar una respuesta.

El estudio comentado anteriormente de Talsma y Woldorff (2005) en el que se observan modulaciones de los potenciales evocados asociados a la localización atendida frente a la no atendida en dos ventanas temporales distintas (i.e., 100-190 ms y 250-500 ms), pone de manifiesto que la interacción entre atención y procesamiento multisensorial puede tener lugar tanto a nivel perceptivo como post-perceptivo.

Investigaciones con neuroimagen como la de Mishra y Gazzaley (2012) muestra actividad neural diferencial en áreas sensoriales primarias (i.e., córtex visual extraestriado, giro temporal medial) así como en áreas heteromodales de integración multisensorial (i.e. giro temporal superior) cuando se atiende a estímulos audiovisuales frente a cuando la atención se centra en una única modalidad sensorial.

Podemos concluir entonces que la atención puede ejercer su influencia tanto a nivel perceptivo como post-perceptivo. Esta distinción debe ser tomada en cuenta a la hora de hacer inferencias sobre cómo atención y procesamiento multisensorial interactúan con el fin de ser lo más precisos posibles en la interpretación de los resultados derivados de los distintos estudios, pudiendo así generar una teoría precisa de lo que ocurre en el cerebro.

3. Congruencia: Manipulando las proporciones

Como comentábamos en el apartado 1.4. al hablar del paradigma de conflicto y las características que se han tenido en cuenta para generar conflicto en el laboratorio, la congruencia es una característica a destacar dado que nuestro sistema cognitivo está preparado para interpretar que las características de los estímulos que pertenecen al mismo evento deben ser congruentes entre ellas. A nivel evolutivo es útil, por ejemplo, que la localización de la que proviene el rugido de un león y el lugar en el que se mueve un arbusto coincidan, puesto que de eso dependerá que se localice rápidamente que un león está preparándose para atacar detrás del arbusto y por lo tanto de ello dependerá la supervivencia. ¿Qué ocurre cuando se da incongruencia entre la información que

recibimos? En algunas ocasiones el sistema cognitivo da un aviso como cuando nos sentimos mareados si la información visual y vestibular no coinciden, o cuando detectamos que el sonido y la imagen en la televisión están desfasados. Pero en la mayor parte de los casos el sistema tiende a hacer una sobrecorrección, como cuando estamos en el cine y percibimos que los diálogos provienen de la pantalla de proyección en lugar de percibirlos de los altavoces que están situados en distintos puntos de la sala de cine. Incluso las correcciones que hace el sistema cognitivo pueden darse en sentido opuesto, es decir que por el mero hecho de que dos estímulos se presenten juntos se consideran formando parte del mismo evento, como cuando una tormenta está muy cercana y la aparición del rayo coincide con el sonido producido por el trueno y se asume que el sonido lo ha producido la misma fuente que ha producido el destello.

Estos ejemplos ponen de manifiesto como el estudio de la congruencia nos puede ayudar a comprender el funcionamiento del sistema cognitivo y, particularmente, el procesamiento de información proveniente de distintas modalidades sensoriales.

3.1. Efecto de congruencia: índice de procesamiento multisensorial

Las tareas en las que se manipula la congruencia se caracterizan por presentar una dimensión estimular relevante y una dimensión estimular irrelevante que pueden ser congruentes o incongruentes entre ellas. En el caso de que sean congruentes, ambas darán lugar a la misma representación y respuesta, mientras en el caso de que sean incongruentes la representación y respuesta desencadenadas serán incompatibles generando competición entre ellas.

A modo de ejemplo consideremos la tarea Stroop. En esta tarea se debe nombrar el color en el que una palabra está escrita (i.e., dimensión relevante), pero crucialmente dicha palabra es el nombre de un color (i.e., dimensión irrelevante). Cuando ambas dimensiones son congruentes, por ejemplo AZUL, las dos dimensiones llevan a la misma respuesta, mientras que cuando las dimensiones son incongruentes, por ejemplo AZUL,

leer el nombre de la palabra lleva a una respuesta diferente a nombrar el color de la tinta en que la palabra está escrita. Típicamente lo que ocurre es que las personas somos más rápidas y cometemos menos errores a la hora de responder azul en la condición congruente (AZUL) que en la incongruente (AZUL). La diferencia en la ejecución entre ensayos congruentes e incongruentes se denomina *efecto de congruencia* y refleja el grado de interferencia por parte de la dimensión irrelevante, de manera que cuanto mayor sea la magnitud del efecto de congruencia mayor será la interferencia generada y en consecuencia menor la capacidad de seleccionar adecuadamente los estímulos relevantes de los irrelevantes (v.gr., Botvinick, Cohen y Carter, 2004; Egner y Hirsch 2005).

El efecto de congruencia del ejemplo utilizado se circunscribe a dimensiones que pertenece a la misma modalidad sensorial (i.e., visual), pero del mismo modo el efecto de congruencia puede ser medido entre modalidades diferentes indicando la interferencia que una modalidad sensorial ejerce sobre otra, así como la capacidad del sistema para seleccionar la información que debe ser integrada. De este modo podemos utilizar el efecto de congruencia como un índice de procesamiento multisensorial. Así, algunos de los estudios de procesamiento multisensorial que hemos presentado a lo largo de esta revisión se podrían reformular en términos de efecto de congruencia. Por ejemplo, Choe y colaboradores (1975) manipularon la congruencia temporal (i.e., sincronía) entre estímulos visuales y auditivos presentados en distintas localizaciones espaciales. Encontraron que cuando ambos estímulos eran presentados de manera sincrónica - condición congruente - se percibían proviniendo de la misma localización (i.e., efecto del ventrílocuo), mientras que cuando la presentación era asincrónica - condición incongruente - esta proporción disminuía hasta aproximadamente la mitad. Por lo tanto la diferencia entre ensayos congruentes e incongruentes - efecto de congruencia - está evidenciando la existencia de procesamiento multisensorial.

Otro estudio a tener en cuenta sería el llevado a cabo por Shams y colaboradores (2000) en el que se introdujo una manipulación de la congruencia numérica entre estímulos visuales y auditivos, dando lugar a una condición congruente en la que un estímulo visual era acompañado de un estímulo auditivo, y a una incongruente en la que era acompañado por más de un sonido. Los participantes juzgaban correctamente que se había presentado un solo estímulo visual en la condición congruente, mientras erraban a la hora de identificar el número de estímulos visuales presentados cuando más de un sonido se presentaba. El efecto de congruencia aquí pone de manifiesto la interferencia que un sonido ejerce sobre el procesamiento visual al ser integrado con éste.

3.2. Efecto de proporción de congruencia específico del contexto

El efecto de congruencia refleja la interferencia ejercida por las distintas dimensiones que entran en juego en el procesamiento de la información en un momento dado. Hay que destacar que la magnitud del efecto de congruencia no permanece constante, sino que puede variar en función de factores como, entre otros, la edad o el aprendizaje, como se pone de manifiesto en un estudio llevado a cabo por Davidson, Zacks y Williams (2003) en el que se observa como la magnitud del efecto de congruencia en la tarea Stroop es menor para jóvenes que para mayores, y como con la práctica de la tarea la magnitud del efecto va disminuyendo.

Otro factor que influye en la magnitud del efecto de congruencia es el denominado *proporción de congruencia*. Al hablar de proporción de congruencia nos referimos a la proporción entre ensayos congruentes e incongruentes, hablando de alta proporción de congruencia cuando hay más ensayos congruentes que incongruentes, y de baja proporción de congruencia cuando se da la situación inversa.

Logan y Zbrodoff (1979) en un estudio pionero observaron una disminución en la magnitud del efecto de congruencia Stroop en una condición en la que había alta proporción de estímulos incongruentes frente a la condición compuesta por una alta

50

proporción de estímulos congruentes. Mostrando así por primera vez como la proporción de congruencia puede modular el efecto de congruencia. La explicación dada por los autores es que la proporción de congruencia cambia el modo de seleccionar la información que es relevante en cada condición, de manera que cuando la mayor parte de los estímulos son congruentes se permite la lectura de la palabra, puesto que ayudará a dar la respuesta correcta, mientras cuando la mayor parte son incongruentes se aprende a inhibir la lectura de la palabra y así lograr una ejecución correcta, responsable de la disminución del efecto de congruencia, ya que la interferencia se minimiza.

Este efecto de la proporción de congruencia fue replicado en estudios que utilizaban el mismo paradigma que Logan y Zbrodoff (1979), es decir donde la presentación de las condiciones de alta y baja proporción de congruencia se hacía en distintos bloques (v.gr., Lindsay y Jacoby, 1994), así como en paradigmas que intercalaban la presentación de ambas condiciones en el mismo bloque de ensayos, asociando un par de ítems como altamente congruentes y otro como altamente incongruentes (v.gr., Jacoby, Lindsay y Hessels, 2003). De esta manera se pretendía comprobar si la selección de la estrategia de lectura (i.e., permitirla o inhibirla) se podía modificar ensayo a ensayo, ya que a diferencia de los estudios anteriores, aquí no se podía predecir si el estímulo presentado iba a pertenecer a la condición de alta o baja proporción de congruencia.

El principal problema con este segundo paradigma, es que al asociar pares de ítems concretos con la probabilidad de que fuesen congruentes o incongruentes, el efecto obtenido puede deberse a la mera asociación entre estímulo y respuesta. Es decir, si los pares de palabras asociados a la alta proporción de congruencia son AZUL y NEGRO, y los asociados a la baja proporción de congruencia son AMARILLO y ROJO, los participantes aprenden que cuando aparece la palabra AZUL se debe responder azul y qué cuando aparece la palabra AMARILLO se debe responder rojo.

Con el fin de poder responder a la pregunta de si verdaderamente la selección de la información relevante en cada caso puede verse modulada ensayo a ensayo Crump,

Gong y Milliken (2006) desarrollaron un paradigma en el que los mismos ítems eran presentados en distintos contextos y cada uno de esos contextos estaba asociado con una proporción de congruencia. Dado que los contextos se presentaban entremezclados, se seguía sin poder predecir cuál sería la proporción de congruencia a la que pertenecería el próximo ítem, y como los mismos ítems aparecían en ambos contextos se eliminaba la posibilidad de un efecto debido a la asociación estímulo-respuesta.

Concretamente en el paradigma que denominaron Proporción de Congruencia Específica del Contexto (CPSC por sus siglas en inglés) se presentaba en el centro de una pantalla una palabra en blanco, que era el nombre de un color, y a continuación encima o debajo de ese punto central se presentaba un rectángulo coloreado, que podía ser congruente o incongruente con el contenido de la palabra previa. La localización (i.e., arriba-abajo) era la que identificaba cada uno de los contextos, de manera que cuando arriba se presentaban los rectángulos siendo un 75% de las veces congruentes en color con el color que era indicado por la palabra previa, abajo el 75% de las veces el color del rectángulo era incongruente con el contenido de la palabra, o a la inversa. Los resultados de este estudio mostraron un efecto de congruencia de menor magnitud para el contexto de baja proporción de congruencia frente al contexto de alta proporción de congruencia. Este efecto se generalizó incluso a ítems nuevos - no asociados a ninguna proporción de congruencia - por el mero hecho de presentarse en el contexto de alta o baja proporción de congruencia (v.gr., Crump y Milliken, 2009).

Podemos concluir entonces que el paradigma CSPC demuestra que el sistema cognitivo es sensible a la asociación entre localización y proporción de congruencia y que además es capaz de utilizar dicha asociación de un modo rápido y flexible (i.e., cambiando del ensayo actual al siguiente) para guiar la selección de la información relevante en cada caso.

3.3. Paradigma CSPC y atención

Los resultados obtenidos aplicando el paradigma CSPC se relacionan con la implicación de procesos atencionales inherentes al paradigma. Por un lado, la constitución de dos contextos mediante la asociación de características como la proporción de la congruencia y la localización no puede darse sin atención. Como han demostrado numerosas investigaciones (v.gr., Jiménez y Méndez, 1999; Treisman, 2006; Crump y cols., 2008) las características de los estímulos, que son procesadas a nivel perceptivo, necesitan de la atención para integrarse en un evento único y para poder aprender la relaciones entre ellas. Así, las distintas características del estímulo como color, posición y congruencia con respecto a la palabra previa están siendo atendidas porque se necesita prestar atención a la posición del estímulo para localizarlo, al color para responder y a la congruencia porque facilitará o obstaculizará los mecanismos que para dar la respuesta habían sido puestos en marcha por la palabra previa.

Por otra parte, mecanismos de atención selectiva también están siendo puestos en marcha, ya que una vez que la asociación entre contexto (i.e., localización) y proporción de congruencia se afianza, la localización en la que se presenta un estímulo pone en marcha los procesos de selección de información relevante adecuados a cada contexto. Por ejemplo, filtrar la información de la palabra leída en un contexto de baja proporción de congruencia, puesto que ésta genera competición entre representaciones y respuestas. Por lo tanto, en este contexto de competición mecanismos atencionales guiados por la meta entran en juego para controlar la ejecución (véase Desimone y Duncan, 1995; Becks y Kastner, 2005).

En línea con la idea de que los mecanismos atencionales que subyacen a los efectos encontrados con el paradigma CSPC se ponen en funcionamiento de manera rápida y flexible, King, Korb y Egner (2012) estudiaron qué ocurría con los efectos de proporción de congruencia específicos del contexto cuando se repetía el mismo contexto en ensayos consecutivos frente a cuando se cambiaba de contexto (i.e., transición de contexto).

CONGRUENCIA: MANIPULANDO LAS PROPORCIONES

Concretamente, la hipótesis alternativa que planteaban era que el set de control atencional válido para el contexto de baja proporción de congruencia (alta competición) estaba siendo adoptado como set general para ambas localizaciones. Si este planteamiento fuese correcto se esperaría que los efectos de proporción de congruencia se observasen independientemente del tipo de transición entre contextos, ya que como comentábamos el set de control puesto en marcha es siempre el mismo. Mientras que si verdaderamente lo que ocurre es que cada localización desencadena un set de control diferente los efectos de congruencia encontrados cuando se repite el contexto y cuando se cambia de contexto deberían ser diferentes. Lo que los autores encontraron fue que los efectos de proporción de congruencia específicos de contexto se observaban sólo tras la repetición del mismo tipo de contexto, mientras que desaparecían cuando se daba un cambio de contexto entre ensayos consecutivos. Este hallazgo apoya la idea planteada en estudios anteriores de que el paradigma CSPC refleja cambios “sobre la marcha” en los procesos de selección atencional.

CAPÍTULO 2. MOTIVACIÓN Y OBJETIVOS

El mundo que nos rodea es de naturaleza multisensorial y es por ello que el estudio del procesamiento multisensorial se convierte en un tema central de investigación para comprender como las personas percibimos coherentemente la información del entorno e interactuamos con él.

Como se ha puesto de manifiesto a lo largo de la revisión llevada a cabo en la introducción de esta tesis, el sistema cognitivo humano esta preparado para ser sensible a determinadas características de los estímulos (v.gr., localización, duración), así como a utilizar los propios recursos cognitivos (v.gr., atención, memoria) en beneficio del procesamiento de toda la información que constantemente llega a nuestros sentidos y crear de este modo una realidad coherente.

De entre todas las características relevantes para este tipo de procesamiento aquí vamos a destacar la congruencia y la atención.

Ya hemos comentado la importancia de manipular la congruencia entre los estímulos para así poder comprobar el modo en que interactúan, por ejemplo, revelando cuál es el peso relativo de cada uno de ellos observando cual modula y cual es el modulado (v.gr., efecto del ventrílocuo), manipulando las condiciones de la situación y observando si la interacción varía (v.gr., efecto del ventrílocuo inverso en situaciones de poca definición del objeto visual)... Dichas modulaciones se cuantifican calculando la diferencia en ejecución entre los ensayos congruentes e incongruentes, lo que denominamos efecto de congruencia, cuya magnitud se utiliza como índice de procesamiento multisensorial. En esta línea de investigación nosotros utilizamos un paradigma de conflicto en el que variamos la congruencia entre estímulos audiovisuales.

Además nos interesa conocer como la atención influye en el procesamiento de esos estímulos congruentes e incongruentes. Se ha visto ya que la atención puede ejercer su influencia mediante un procesamiento de abajo-arriba guiado por los estímulos, o bien mediante un procesamiento de arriba-abajo guiado por las metas (véase Talsma y cols., 2010). En esta argumentación sobre el papel de la atención se han llevado a cabo

MOTIVACIÓN Y OBJETIVOS

muchos estudios, pero ninguno de ellos ha manipulado la proporción de congruencia entre las características de los estímulos. Nosotros utilizando el paradigma propuesto por Crump y colaboradores (2006) manipulamos la proporción de congruencia audiovisual de dos modos diferentes y asociamos cada uno de ellos con una localización distinta, creando así lo que denominaremos contextos de congruencia. Como ya se ha discutido anteriormente en esta manipulación del contexto de congruencia están implicados mecanismos de atención selectiva que guían el procesamiento de los estímulos que son relevantes en cada una de las localizaciones.

La utilización de este paradigma tiene dos ventajas principales, por un lado que al manipular la congruencia tenemos un índice directo de procesamiento multisensorial - efecto de congruencia -. Por otro lado que la manipulación de la proporción de congruencia nos permite comprobar si la atención modula el efecto de congruencia (procesamiento multisensorial) en función del contexto en el que se presentan los estímulos.

Con el objetivo de estudiar el alcance del efecto de la atención en el procesamiento audiovisual tendremos en cuenta también el nivel de procesamiento al que se dan tanto el efecto de congruencia (procesamiento multisensorial) como el efecto de congruencia específico del contexto (atención). Para ello utilizaremos técnicas psicofísicas como la teoría de detección de señales, técnicas de electroencefalografía y técnicas de estimación de fuentes neurales.

A continuación se describen las series experimentales llevadas a cabo con el objetivo general de profundizar en el conocimiento del papel de la atención en el procesamiento de estímulos audiovisuales. En cada una de las series se describirán los objetivos específicos, así como el paradigma y el tipo de técnica utilizada para el registro y análisis de los datos.

Serie Experimental I: Audiovisual interactions depend on context of congruence

La Serie Experimental I está compuesta por cuatro experimentos comportamentales.

El Experimento 1 tiene como objetivo encontrar una manipulación de congruencia de estímulos audiovisuales simples que sea susceptible de mostrar un efecto multisensorial. Concretamente se presenta en el centro de la pantalla un círculo blanco que puede durar 40 ó 120 ms, y que puede ir acompañado de un ruido blanco de igual (i.e., congruente) o diferente (i.e., incongruente) duración. La tarea consiste en discriminar si el círculo tiene una duración corta o larga. La hipótesis que planteamos es que el sonido sesga el procesamiento del estímulo visual, esto es, se muestra una mejor ejecución en los ensayos congruentes y un deterioro en los incongruentes si comparamos ambos ensayos con aquellos en los que sólo se responde a estimulación visual. Esperamos que el sonido tenga tanto peso porque estamos ante una tarea de duración y, como hemos comentado anteriormente, el sentido de la audición tiene mayor resolución en el ámbito temporal que el sentido de la visión.

En el Experimento 2 se añade a esta tarea de discriminación la manipulación del contexto de congruencia, con el objetivo de comprobar si el efecto de congruencia se ve modulado por la mayor o menor presencia de estímulos congruentes. Específicamente la tarea de discriminación audiovisual se realiza primero en un contexto en el que el 75% de los ensayos son congruentes y tras varios bloques se cambia a un contexto en el que el 75% de los ensayos son incongruentes, o viceversa. Nosotros esperamos que el sistema se haga sensible a esta proporción y que el efecto de congruencia sea menor en el caso del contexto con baja proporción de ensayos congruentes. Esta hipótesis se basa en la idea de que cuando hay muchos estímulos incongruentes la aproximación de selección de estímulos adecuada pasa por filtrar la información auditiva, información que en la mayor parte de las ocasiones lleva a una respuesta errónea. Mientras en el contexto con alta proporción de ensayos congruentes resulta útil prestar atención a la información visual y

MOTIVACIÓN Y OBJETIVOS

auditiva para responder correctamente en los ensayos congruentes, pero este proceso de selección también lleva a error en los pocos ensayos que son incongruentes, siendo así responsable de la mayor magnitud del efecto de congruencia (i.e., ensayos congruentes - ensayos incongruentes).

Los dos experimentos restantes - Experimento 3 y Experimento 4 - pretenden comprobar si el cambio en la atención selectiva puede ocurrir de manera rápida y flexible. Para ello se utiliza el mismo procedimiento que en el Experimento 2, pero esta vez cada contexto se presenta en una localización diferente de la pantalla (i.e., encima o debajo del punto central), dando así la posibilidad de intercalar aleatoriamente la presentación de ambos contextos. La hipótesis que se plantea es que nuevamente la magnitud del efecto de congruencia será menor en el contexto de baja proporción de ensayos congruentes. Esto estaría indicando que los mecanismos de selección de información se despliegan de manera rápida y flexible en función de la localización en la que aparezca el estímulo.

En los Experimentos 2, 3 y 4 de esta serie experimental tras finalizar la tarea se recogen mediante cuestionario medidas subjetivas que indican el grado de conocimiento que los participantes tienen sobre la manipulación de congruencia. Dado que estudios previos habían demostrado que el efecto de congruencia específico del contexto ocurre en ausencia de consciencia por parte de los participantes (v.gr., Crump y cols., 2006; Heinemann, Kunde y Kiesel, 2009) nosotros esperamos replicar ese mismo resultado.

Además con el objetivo de explorar el nivel de procesamiento al que está ocurriendo tanto el procesamiento multisensorial como el efecto de CSPC, se aplica un análisis basado en la teoría de detección de señales.

Con la finalidad de poder investigar en profundidad el tema de los niveles de procesamiento en la Serie Experimental II se añade el registro de actividad eléctrica cerebral durante la realización de la tarea.

Serie Experimental II: Electrophysiological evidence of audiovisual interactions under different contexts of stimulus congruence

La Serie Experimental II tiene dos objetivos principales. Por un lado estudiar si el efecto de congruencia audiovisual observado en la Serie Experimental I implica integración de los estímulos, entendida como una diferencia cualitativa entre la actividad eléctrica asociada al procesamiento de los estímulos audiovisuales frente a la suma de las actividades asociadas a cada uno de los estímulos por separado (i.e., $AV \neq A+V$). Por otro lado conocer el nivel de procesamiento en el que la manipulación del contexto de congruencia influye sobre el efecto de congruencia audiovisual.

El primero de los objetivos se pretende evaluar mediante la aplicación de una tarea oddball, en la que los estímulos auditivos y visuales (i.e., círculo y ruido blanco) se presentaban en las dos localizaciones del espacio, pero en ausencia de manipulación de la proporción de congruencia. Los estímulos podían ser presentados solos (i.e., condición unimodal; A y V) o en sincronía (i.e., condición multisensorial; AV), permitiendo así la comparación de la actividad asociada a la suma de las condiciones unimodales con la condición multisensorial y responder al objetivo propuesto. Para eliminar interferencias producidas por la selección y ejecución de respuesta, los estímulos objetivos, es decir los que requieren una respuesta, son un triángulo y un tono puro.

El segundo de los objetivos se evaluó utilizando el mismo paradigma que en la Serie Experimental I. La hipótesis que se plantea es que el efecto de la atención ocurre en una ventana temporal temprana, de manera que se apoye la idea de que la atención modula la integración de los estímulos audiovisuales y no sólo tiene un papel a la hora de seleccionar la respuesta.

El análisis de estas dos tareas se ha llevado a cabo mediante un procedimiento multivariado denominado *Electrical neuroimaging* (véase Murray, Brunet y Michel, 2008). A continuación se explica el porqué de la elección de esta técnica así como una breve descripción de la misma.

MOTIVACIÓN Y OBJETIVOS

Basándonos en los principios biofísicos sabemos que la actividad eléctrica generada por las neuronas se propaga en todas direcciones creando el campo de voltaje que puede ser registrado desde el cuero cabelludo. Esto tiene como consecuencia que un electrodo individual registre actividad de las diferentes fuentes activas en el cerebro en un momento dado, las cuales debido a la propagación pueden estar bastante distantes del electrodo concreto. Es decir, no hay una relación directa entre la localización del electrodo y la posición de las neuronas activas. Es por ello que esta técnica toma en consideración la totalidad de los electrodos que constituyen el montaje.

Tomando la actividad registrada por todos los electrodos en el cuero cabelludo podemos reconstruir para cada instante temporal el mapa de voltaje correspondiente. Dicho mapa de voltaje se caracteriza por su potencia y por su topografía. La potencia indica la cantidad de fuentes que están activas simultáneamente, mientras que la topografía se relaciona con la distribución de las fuentes. De hecho, acorde con las leyes de la física, siempre que un mapa topográfico cambia, la distribución de las fuentes activas también cambia (véase Vaughan, 1982; Lehmann, 1987).

La técnica *electrical neuroimaging* permite analizar estos dos parámetros de manera independiente mostrando, por ejemplo, si hay diferencias en la topografía entre dos condiciones, si la topografía es la misma pero la potencia de las fuentes es diferente entre condiciones, o incluso si tanto potencia como topografía varían de una condición a otra.

La medición de estos dos parámetros se hace mediante un índice de la Potencia de Campo Global - *Global Field Power (GFP)* - en el caso de la potencia, y mediante un índice de Disimilitud global - *Global Dissimilarity (Diss)* - en el caso de la topografía (véase Lehmann y Skrandies, 1980). En concreto, el GFP es un valor que indica la desviación estándar del conjunto de valores de los electrodos en un momento temporal concreto, mientras la Diss se define como la diferencia entre dos campos eléctricos y se calcula como la raíz cuadrada de la media de las diferencias al cuadrado entre los potenciales

medidos en cada electrodo, que han sido previamente normalizados al dividir su valor por el GFP.

Por último, a partir de la información obtenida a nivel de potencia y de topografía se lleva a cabo una estimación de la localización de las fuentes neurales que podrían estar generando la actividad registrada en el cuero cabelludo. Para ello utilizamos un modelo de solución inversa denominado *Local Autoregressive aveRAge (LAURA)* (véase Michel, Murray, Lantz, González, Spinelli, 2004) que realiza una reconstrucción tridimensional de la actividad eléctrica cerebral incorporando las limitaciones propias de la biofísica del cerebro.

Esta serie experimental consta de dos experimentos, sin embargo, en el texto principal de la tesis sólo se presenta el segundo de ellos - Experimento 6 -, porque los resultados del primero - Experimento 5 - aunque similares (ver Anexo I), no pueden ser tomados como fiables debido a un problema con el sistema de registro. Sin embargo, es relevante esta aclaración, dado que el diseño del Experimento 6 es sucesor del diseño y resultados del primero de los experimentos de esta serie.

Cuatro son las mejoras añadidas al diseño del Experimento 6 en relación al primero que se realizó. En primer lugar, se modificó la duración de los estímulos, pasando de durar 40 y 80 ms, como en los estudios comportamentales, a durar 100 y 180 ms. La idea que subyacía a este cambio era intentar disociar los efectos de integración multisensorial, que se pueden comenzar a observar alrededor de los 50 ms, de efectos debidos a la disparidad en la duración entre los estímulos auditivos y visuales, que debido a la duración de nuestros estímulos podrían observarse también en la misma ventana temporal. En segundo lugar, los bloques estaban constituidos por un mayor número de ensayos, lo que permitiría analizar los efectos de transición de contexto en línea con los resultados que en 2012 presentaron King y colaboradores. En tercer lugar, se aumentó la tasa de muestreo en el registro de la actividad electroencefalográfica, pasando de 250Hz a 1000Hz. Dado que los efectos tempranos se caracterizan por ser temporalmente breves,

MOTIVACIÓN Y OBJETIVOS

la adquisición de pocas muestras temporales podría conllevar la pérdida del efecto debido a la modificación provocada en la señal registrada. Por último, tras comprobar que la técnica de *electrical neuroimaging* era sensible a los efectos de proporción de congruencia específicos de contexto, el siguiente paso fue añadir la tarea oddball con el objetivo de tener una línea de base de la integración de los estímulos audiovisuales previa a la manipulación del contexto de congruencia.

CAPÍTULO 3. SERIE EXPERIMENTAL I. AUDIOVISUAL INTERACTIONS DEPEND ON CONTEXT OF CONGRUENCE

El contenido de este capítulo ha sido publicado como Sarmiento, B.R., Shore, D.I., Milliken, B. y Sanabria, D. (2012). Audiovisual interactions depend on context of congruency. *Attention, Perception & Psychophysics*, 74, 563-574.

1. Abstract

In this study, we addressed how the particular context of stimulus congruence influences audiovisual interactions. We combined an audiovisual congruence task with a proportion of congruence manipulation. In Experiment 1, we demonstrated that the perceived duration of a visual stimulus was modulated by the actual duration of a synchronously-presented auditory stimulus. In the following experiments we demonstrated that this crossmodal congruence effect was modulated by the proportion of congruent trials between (Experiment 2) and within blocks (Experiment 4). In particular, the crossmodal congruence effect was reduced in the context with a high proportion of incongruent trials. This effect was attributed to changes in participants' control set as a function of the context of congruence, with greater control applied in the context where the majority of the trials were incongruent. These data contribute to the ongoing debate concerning crossmodal interactions and attentional processes. In sum, context can provide a powerful cue for selective attention to modulate the interaction between stimuli from different sensory modalities

2. Introduction

Research in the field of cognitive control has been tightly linked to the study of stimulus congruence. The current research aimed to study the interplay between audiovisual interactions and attentional control processes driven by changes in context, i.e., the relative proportion of congruent and incongruent trials.

In a typical congruence task (e.g., Stroop), two stimulus dimensions, one task-relevant and one task-irrelevant, can be congruent (triggering the same response), or incongruent (triggering incompatible responses). For example, in the Stroop task, the colour word and the colour of the ink in which the word is written can be congruent (e.g., red written in red) or incongruent (e.g., red written in blue). The difference in performance

INTRODUCTION

between congruent and incongruent trials (i.e., faster RT and higher response accuracy on congruent trials) is called *congruence effect*. It is argued (e.g., Botvinik, Cohen & Carter, 2004; Egner & Hirsch, 2005) that the size of the congruence effect reflects the level of control that the participant exerts to avoid the interference of irrelevant information while making a fast and accurate response. By this view, the larger the size of the congruence effect, the more interference occurred, and thus, less control was exerted.

A number of studies have shown that the congruence effect can vary as a function of the proportion of congruent trials. In a seminal paper, Logan and Zbrodoff (1979) manipulated the proportion of congruent and incongruent trials presented in a Stroop task, finding a decreased congruence effect in those blocks of trials with a higher proportion of incongruent items. These authors (see also Lindsay & Jacoby, 1994) concluded that this proportion-congruent effect was due to changes in word-reading strategies in response to changes in the likelihood of congruence.

Crump, Gong and Milliken (2006), using the same rationale as the abovementioned studies, manipulated the proportion of congruent items presented in a consecutive trial variant of the Stroop task. A colour word prime was presented in white at fixation and was followed by a coloured rectangle target that appeared above or below fixation. The colour of the target either matched (i.e., was congruent) or mismatched (i.e., was incongruent) the meaning of the preceding colour word prime. The participants' task was to identify the colour of the rectangle. Unlike previous experiments, the coloured rectangle targets were presented in either of two different contexts (above or below a central fixation point) mixed at random within blocks of trials. One of these contexts was associated with a high proportion of congruent trials (75%) and the other was associated with a low proportion of congruent trials (25%). Participants in Crump et al.'s study could not predict which proportion-congruent context the current trial belonged to until the onset of the coloured target itself. Hence, participants could not take advantage of preparatory strategies prior to the onset of the prime word that would facilitate or inhibit word-reading in accord with the

68

proportion of congruence. Nonetheless, the results of Crump et al. showed that the proportion-congruent context, even when manipulated randomly at target onset within a block of trials, modulated the Stroop interference in accordance with previous studies: a significantly reduced congruence effect on the low proportion-congruent context compared to the high proportion-congruent context. They proposed that these context-sensitive control effects reflected learning of associations between context (e.g., location) and likelihood of congruence, giving rise to context-specific proportion congruent (CSPC) effects.

Crump et al. (2006) suggested that attentional processes were responsible for the CSPC effects obtained in their study. On one hand, selective attention was necessary to associate a proportion of congruence with a particular context (see Crump, Vaquero & Milliken 2008, Experiment 3). On the other hand, this association implied that a particular context triggered attentional adjustments to select the suitable incoming information (Crump & Milliken, 2009).

The critical question addressed in our study is whether the CSPC effects can be also obtained when congruence refers to a stimulus dimension shared between inputs from two sensory modalities. The manipulation of congruence using multisensory stimuli allows the testing of mechanisms of crossmodal interaction (see Calvert, Spence, & Stein, 2004). In particular, when inputs from two sensory modalities are incongruent, performance can reveal whether an input from one sense is able to affect the perception of inputs in the other sensory modality. In effect, both behavioural (e.g., Sanabria, Spence & Soto-Faraco, 2007) and neuroimaging (e.g., Alink, Singer & Muckli, 2008; Watkins, Shams, Josephs, & Rees, 2007) research has shown that crossmodal congruence effects can reflect interactions at perceptual level of information processing prior to attentional deployment, in contrast with unimodal congruence effects, such as the Stroop effect, that occur at higher post-perceptual levels of information processing.

INTRODUCTION

Relevant here is the study by Shore and Simic (2006) who showed a modulation of the visuotactile congruence effect by the relative proportion of congruent and incongruent trials. The authors manipulated the proportion of congruent and incongruent trials between blocks of trials, obtaining a reduced congruence effect for errors in blocks of trials with a 25% of congruent trials with respect to blocks of trials with a 75% of congruent trials. However, this effect was only obtained when the visual distractor was presented 100 ms prior to the tactile target. The authors concluded that visuotactile integration was immune to top-down influences, and, if any, these influences could be exerted only when the two sensory inputs were presented in asynchrony. Therefore, it remains unsolved whether the CSPC effects could be obtained in the crossmodal domain when the two sensory inputs are presented in synchrony (i.e., maximising the crossmodal interaction process) and, moreover, whether the CSPC effects can be shown when two different context of congruence are manipulated within the same block of trials like in Crump et al.'s (2006) study.

The congruence manipulation in our study would produce more accurate responses on congruent trials and more errors on incongruent trials with respect to a baseline unimodal visual condition. A significant difference in accuracy was also predicted between congruent and incongruent trials. Crucially, in line with Crump et al.'s argument about the implication of attentional processes in CSPC effects, we expected a larger congruence effect in the high proportion-congruent context than in the low proportion-congruent context supporting the idea that audiovisual interactions depend on the deployment of attentional control, in line with relevant empirical evidence from behavioural (e.g., Sanabria, Soto-Faraco & Spence, 2007), ERP (e.g., Talsma, Doty & Woldorff, 2007) and neuroimaging studies (e.g., Fairhall & Macaluso, 2009) that have demonstrated these top-down influences.

In Experiment 1, we developed an audiovisual task that resulted in a congruence effect whereby the perceived duration of a visual stimulus was modulated by the

synchronous presentation of an auditory stimulus of the same (congruent) or different duration (incongruent) (see Klink, Montijn & Wezel, 2011, for a somewhat similar procedure). The data showed more accurate performance on congruent than on incongruent trials, revealing a congruence effect similar to that reported in previous studies (e.g., Fairhall & Macaluso, 2009). Signal detection theory analyses revealed that auditory inputs influenced visual perception at a perceptual level of processing (see Sanabria et al., 2007; Watkins, Shams, Josephs, & Rees, 2007). In Experiments 2-4, we associated the audiovisual stimuli with two different contexts, with regard to the likelihood of congruence. The particular context of congruence was manipulated between blocks of trials (Experiment 2) and within blocks of trials (Experiment 3 and 4). On the basis of what was found by Crump and colleagues (2006, 2008 & 2009), we expected a reduced crossmodal congruence effect in the context with more incongruent trials than in the context with a higher proportion of congruent trials.

The combination of Crump and colleagues' paradigm (2006) and crossmodal congruence constitutes a novel approach, which allows an innovative way of studying the role of attention in crossmodal interaction. Our findings would then shed new light on the interplay between attentional processes and crossmodal interactions (see Talsma, Senkowski, Soto-Faraco & Woldorff, 2010, for a review).

3. Experiment 1

3.1. Method

3.1.1. Participants

The participants were 19 undergraduate students (13 females; age range: 17-43 years old; mean age of 19 years old) who received course credits in exchange of their participation. All of the participants in this study reported normal hearing and normal or corrected-to-normal vision, and gave their informed consent to participate in the study

EXPERIMENT 1

conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

3.1.2. Apparatus and stimuli

The experiment was conducted on an Intel Core 2 Duo PC with a 17-in. LCD monitor. E-Prime software was used for stimulus presentation and response collection (Schneider, Eschman, & Zuccolotto, 2002).

The visual stimuli consisted of a white cross that served as the fixation point and a target white circle (3.01° in diameter). Two loudspeakers, positioned on each side of the computer screen, were used to present the auditory stimulus, which consisted of a white noise burst (60 dB[A] measured at ear level).

3.1.3. Procedure

Participants sat in a comfortable chair at approximately 57 cm from the computer monitor in a dark room. They were asked to discriminate the duration (short or long) of a central white circle displayed below (4.52° from the centre of the circle) the fixation point for either 40 ms or 120 ms, while ignoring the synchronous presentation of a white noise burst that could last for 40 ms or 120 ms. This manipulation gave rise to congruent trials (the visual stimulus was of the same duration as the auditory stimulus) and incongruent trials (the visual stimulus was of a different duration than the auditory stimulus).

At the beginning of each trial, participants were presented with a fixation cross displayed in white against a black background for a random duration between 350 and 1250 ms that remained on and steady for the whole trial. Then, the circle target was presented for either 40 ms or 120 ms accompanied by the white noise burst that could be congruent or incongruent in duration. The short and long stimuli had the same probability of appearance. This was also true for congruent and incongruent trials. Accuracy, rather than response speed, was stressed, so participants had no response time pressure. Half of the participants were told to press the “C” key if the circle was short in duration and the “N” key if the circle was long in duration. The reverse stimulus-response mapping was

used for the remaining participants. Feedback of response accuracy was provided to the participant. The next trial began 1000 ms after the feedback. (See Figure 2).

The experiment began with a unimodal discrimination task (the visual target stimuli were presented in the absence of auditory distractors) in which participants completed 8 unimodal practice trials, followed by a block of 40 (20 short & 20 long) unimodal experimental trials. The unimodal block was followed by the crossmodal discrimination task in which participants completed 8 crossmodal practice trials and 4 blocks of 40 crossmodal trials each.

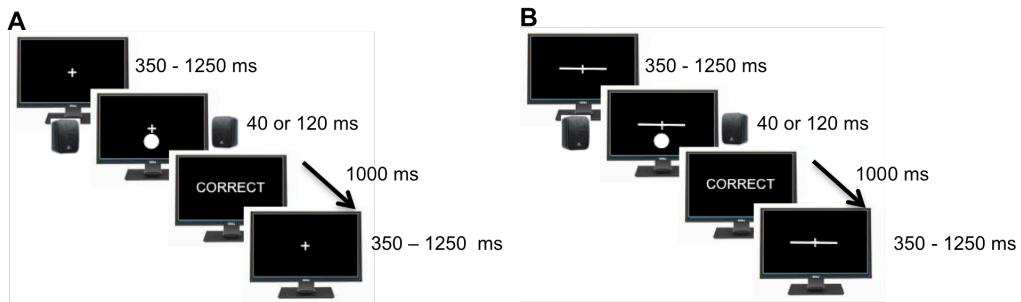


Figure 2. Schematic illustration of the set up used in Experiments 1-3 (A) and Experiment 4 (B).

3.2. Results

Data from one participant were removed from the analysis because her performance did not reach 60% accuracy on the baseline condition. The remaining 18 subjects (12 females; age range: 17-43 years old; mean age of 19 years old) were included in the complete data analysis.

Mean response accuracy¹ for each participant and condition were analysed using pair-wise two-tailed t-tests.

¹ In contrast to the majority of studies using congruency tasks that used RT, accuracy was our main dependent variable. We believe that accuracy was the more suitable dependent variable for the following reasons: 1) temporal pressure in our study would have resulted in even a greater number of errors, which would have prevented from obtaining enough correct responses in order to analyze the RT data. 2) the focus of interest here was response precision, not response speed. 3) since participants in the task used in our study committed errors even without temporal pressure, accuracy was considered as a very informative variable of audiovisual interactions, given that those errors

EXPERIMENT 1

There was a significant difference between congruent and incongruent trials, $t(17) = 10.98$, $p < .001$ (see Figure 3). Responses on congruent trials (83% correct) were more accurate than responses on incongruent trials (45% correct).

The comparison between each type of trial and the baseline (unimodal condition) reached statistical significance, $t(17) = 3.17$, $p = .005$ and $t(17) = 8.78$, $p < .001$ for congruent and incongruent, respectively.

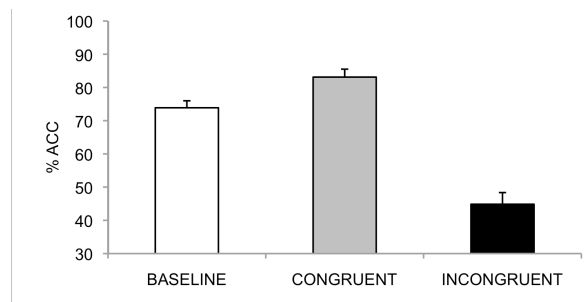


Figure 3. Percentage of correct responses for Unimodal and Crossmodal (congruent and incongruent) trials in Experiment 1. Error bars represent the standard error of the mean across participants.

Sensitivity (d') and criterion (c) indexes were assessed in baseline, short auditory and long auditory conditions for each participant based on the proportions of hits (correct short visual target detections in short target trials) and false alarms (incorrect short visual responses in long target trials).

Hypothesis-driven pair-wise two-tailed t -tests revealed differences in d' between baseline unimodal (1.51) condition and both crossmodal conditions (short (0.89) and long (0.89) auditory), $t(17) = 3.55$, $p = .002$ and $t(17) = 3.19$, $p = .005$, respectively. Revealing a higher sensitivity in the unimodal condition. The difference in d' between the two crossmodal conditions was not significant, $t < 1$. The difference in c between the baseline condition (0.32) and the two crossmodal conditions (0.69 for short and 0.52 for long), and

were not committed due to response pressure.

Note also that other studies using tasks similar to ours have also used accuracy as their main dependent variable (e.g., Klink, Montijn & Wezel, 2011; Shams, Kamitani & Shimojo, 2002) In any case, the analyses on the RT data from Experiments 1-4 did not revealed any statistical significant effect.

the difference between the two crossmodal conditions themselves, were significant; all $p < .003$. The data revealed a bias toward responding short duration for the circle when it was presented with a short auditory stimulus, and a bias toward responding long when the duration of the auditory stimulus was long.

3.3. Discussion

Experiment 1 showed that participants committed more errors on incongruent trials than on congruent trials, in the visual temporal discrimination task. This result, in accordance with previous studies (e.g., Shams, Kamitani & Shimojo, 2000, 2002; Klink, Montijn & Wezel, 2011), suggests that the auditory stimulus altered the perception of the visual stimulus when the two synchronously presented stimuli were incongruent in duration.

Signal detection theory (SDT) analyses suggest that this modulation occurs at both perceptual (d') and decisional (c) levels of information processing (see Sanabria, Spence & Soto-Faraco, 2007). Crucially, the effect in d' on crossmodal trials (as compared to the baseline unimodal condition) did not depend on whether the auditory stimulus was short or long in duration.

In Experiment 2, we investigated whether the congruence effect found in Experiment 1 could be modulated by the likelihood of congruent trials. To this end, the multisensory stimuli were presented in two contexts with different frequencies of congruent trials. SDT analyses were conducted for each context of congruence in order to study at which level of processing the modulation by context would occur.

4. Experiment 2

4.1. Method

4.1.1. Participants

The participants were 32 undergraduate students (20 women; age range: 17-39 years old; mean age 19 years old) who volunteered for course credit.

4.1.2. Apparatus, stimuli and procedure

These were the same as in Experiment 1 except for the following: In Experiment 2, we included two different congruence contexts. For half of the experimental blocks (High congruence context) 80% of the trials were congruent, and the remaining 20% of the trials were incongruent. The reverse was true for the remaining half of the blocks (Low congruence context).

Participants performed 8 blocks of 40 trials, 4 in each of the High and Low congruence context conditions. Half of the participants performed 4 blocks of trials in the high congruence context condition followed by the 4 blocks of trials in the low congruence context condition. The reverse order of presentation of the congruence context conditions was used for the remaining half of the participants. After the first 4 blocks of trials, participants were informed that they had completed half of the crossmodal discrimination task. At the end of the experiment, participants completed a questionnaire to evaluate whether they were aware of the congruence manipulation. They were asked whether they perceived any difference between the two halves of the crossmodal discrimination task and what that difference might be. They were also asked regarding the reliability of the feedback information (i.e., whether they believed that the information provided by the feedback was in accordance with their response).

4.2. Results

4.2.1. Accuracy measures

Data from eight participants were removed from the analysis because their performance did not reach 60% accuracy in the baseline condition. We set this criterion to make sure that participants understood the task at hand. The remaining 24 subjects (14 females; age range: 17-39 years old; mean age of 19 years old) were included in the complete data analysis.

Mean accuracy for each participant and condition were submitted to a 2 Congruence (congruent, incongruent) by 2 Proportion congruent (high, low) ANOVA, which revealed a significant main effect of Congruence, $F(1,23) = 64.88$, $MSE = 2.59$, $p < .001$, $h^2_p = .74$. Responses on congruent trials (80% correct) were more accurate than responses on incongruent trials (47% correct). The main effect of Proportion congruent was not significant, $F < 1$.

Crucially, the interaction between Congruence and Proportion congruent was significant, $F(1,23) = 42.96$, $MSE = .57$, $p < .001$, $h^2_p = .65$. The size of the congruence effect (% correct on congruent trials - % correct on incongruent trials) was smaller for the low proportion-congruent context (17%) than for the high proportion-congruent context (48%). A priori comparisons revealed that this difference was due to lower accuracy on congruent trials (15.3%), $F(1,23) = 23.72$, $MSE = .28$, $p < .001$, and higher accuracy on incongruent trials (15.5%), $F(1,23) = 12.35$, $MSE = .29$, $p < .001$, in the low proportion-congruent context with respect to the high proportion-congruent context (see Table 1). Participants performed more accurately on congruent trials than on incongruent trials in both proportion-congruent conditions, $F(1,23) = 95.53$, $MSE = 2.8$, $p < .001$ and $F(1,23) = 15.21$, $MSE = .36$, $p < .001$, for the high and low proportion of congruence context, respectively.

EXPERIMENT 2

Experiment	Proportion Congruent	Congruent		Incongruent	
		M	SE	M	SE
2	High	87.60	1.83	39.32	4.08
	Low	72.27	3.54	54.85	3.46
3	High	82.94	2.24	53.65	3.40
	Low	81.59	2.78	52.65	3.41
4	High	77.57	2.21	53.82	2.64
	Low	77.03	2.38	57.18	2.64

Table 1. Mean correct responses (in percentage) and standard errors, for Experiments 2-4

To investigate whether the CSPC effect depend on the learned association between the context and the likelihood of congruence as suggested by Crump et al. (2006), we compared participants' performance in Experiment 2 between the first two blocks and the last two blocks of each context. Mean accuracy for each participant and condition were submitted to a 2 Congruence (congruent, incongruent) by 2 Proportion congruent (high, low) by 2 Block (first half, second half) ANOVA. The first ten trials of the first block were considered part of the training phase and were excluded from the analysis.

Critically, the three-way interaction between Congruence, Proportion congruent and Block was significant $F(1,23) = 10.38$, $MSE = .099$, $p < .005$, $h^2_p = .31$ (see Figure 4). While accuracy in incongruent trials was significantly higher in the low proportion-congruent context with respect to the high proportion-congruent context in the second half of each context, $F(1,23) = 19.47$, $MSE = .57$, $p < .001$, that difference did not reached statistical significance in the first half of each context, $F(1,23) = 1.68$, $MSE = .58$, $p = .21$.

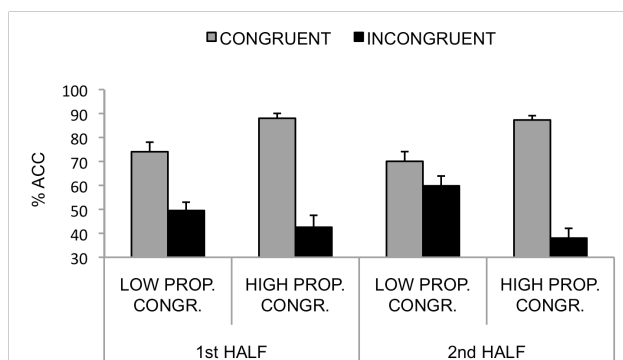


Figure 4. Percentage of correct responses as a function of Congruence, Proportion congruent and Block in Experiment 2. Error bars represent the standard error of the mean across participants.

Short and long auditory conditions were pooled for each context of congruence (high and low) to obtain sensitivity (d') and criterion (c) indexes. Pair-wise two-tailed t -tests showed non-significant differences on d' and c between high and low proportion-congruent context ($t < 1$, $t(23) = 1.24$, $p = 0.22$ respectively).

4.2.2. Questionnaire measures

Regarding the congruence manipulation, only half of the participants (46%) noticed a difference between the two proportion-congruent conditions, although none of them appeared to be aware of the proportion manipulation (see Table 2). Instead, those who noted a difference between the two halves of the experimental session offered comments like “the first part was easier due to tiredness” or “I made fewer mistakes on the second part of the experiment presumably due to practice”.

Experiment	Difference	No difference
2	45.83	54.17
3	16.67	83.33
4	16,67	83,33

Table 2. Percentage of participants that noticed or not any difference between the two contexts of congruence in Experiments 2-4.

To determine whether performance of participants who noticed a difference between the two contexts of congruence significantly differ from performance of participants who did not, a 2 Congruence (congruent, incongruent) by 2 Proportion congruent (high, low) mixed design ANOVA with Context difference (difference, no difference) as a between-participants factor was conducted.

The three-way interaction of Congruence x Proportion congruent x Context difference was significant $F(1,22) = 4.66$, $MSE = .053$, $p < .05$, $h^2_p = .17$ (see Figure 5). Although both groups of participants showed a significant interaction between Congruence and

EXPERIMENT 2

Proportion congruent [$F(1,10) = 9.74$, $MSE = .117$, $p < .002$, $h^2_p = .49$, $F(1,12) = 45.92$, $MSE = .508$, $p < .001$, $h^2_p = .79$ for difference and no difference group, respectively], the difference in the size of the congruence effect between contexts was larger for participants who did not notice any difference between the two contexts (39,54%) than for participants who notice a difference (20,60%). In both cases, the CSPC effects were driven by changes in accuracy on incongruent trials, more pronounced in the group of participants who did not notice a difference between the two contexts of congruence (8% and 21%, for participants who noticed a difference and those who did not, respectively).

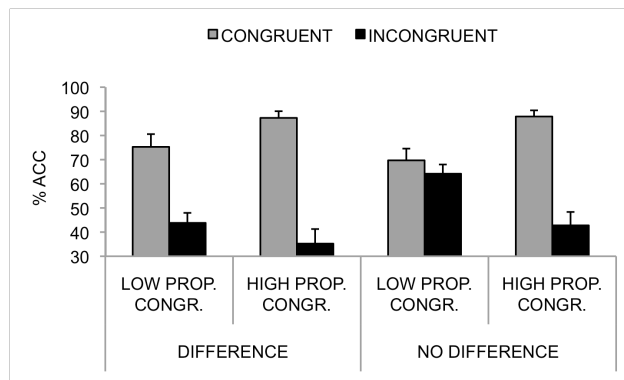


Figure 5. Percentage of correct responses as a function of Congruence, Proportion congruent and Context difference in Experiment 2. Error bars represent the standard error of the mean across participants.

The majority of the participants (83%) considered the feedback information unreliable (see Table 3). At the end of the experiment, they would report that the “incorrect response” feedback display appearing at the end of some trials (presumably incongruent trials) was not in accordance with their performance in those trials. For instance, a participant in a “visual short (40 ms)-auditory long (120 ms)” trial would respond “long”, getting an “incorrect response” feedback display. However, he/she would believe that his/her response was correct. Therefore, the information provided by the feedback in that case was considered unreliable by the participant. This points to the idea that the crossmodal effect reported here can be considered as an audiovisual illusion whereby the perceived

duration of the visual event was driven by the actual duration of the auditory stimulus (cf. Soto-Faraco, Spence, Kingstone, 2004).

Experiment	<i>Unreliable</i>	<i>Reliable</i>
2	83.33	16.67
3	79.17	20.83
4	66.67	33.33

Table 3. *Percentage of participants who considered feedback information as reliable or unreliable in Experiments 2-4.*

4.3. Discussion

Three main results were obtained in Experiment 2. First, we replicated the congruence effect described in Experiment 1. Second, and more importantly, the congruence effect was modulated by the congruence context. Specifically, the congruence effect was reduced in the low proportion-congruent context relative to the high proportion-congruent context. This contrasts with the results reported by Shore & Simic (2006) who failed to show a modulation of the visuotactile congruence effect by the proportion of congruence manipulation when the two inputs were presented in synchrony. Third, the modulation of the congruence effect was emphasized on the second half of each context, presumably due to the learning of the association between the context of congruence and the proportion of congruent trials. SDT results did not show any effect in either the d' or c index.

The results of the questionnaire used to assess whether participants were aware of the proportion-congruent manipulation appeared to indicate no evidence of participants' explicit knowledge about the congruence manipulation. What's more, participants who reported to have noticed a difference between the two contexts of congruence (whatever that difference was) showed a reduced CSPC effects with respect to those participants who did not notice a difference. This suggests that the different performance for the two

EXPERIMENT 3

contexts was not due to voluntary control, but rather to some form of cognitive control that was involuntarily cued by the context.

Crump and colleagues (2006, 2008, 2009) obtained their results using a within blocks congruence manipulation, while data from Experiment 2 came from a between blocks congruence manipulation. We conducted Experiment 3 to test whether the same results could be obtained when the proportion of congruence was manipulated within blocks of trials.

5. Experiment 3

5.1. Method

5.1.1. Participants

The participants were 36 undergraduate students (25 women, age range: 18-30 years old, mean age of 20 years old) who received course credits in exchange for their participation.

5.1.2. Apparatus, stimuli and procedure

These were the same as in Experiment 2 with the exception that two congruence contexts were presented within the same block of trials. The two halves of the computer screen (top/bottom) defined the two congruence contexts. As such, the white circle (3.01° in diameter) was presented below or above a fixation point (4.52° from the centre of the circle). The fixation point was presented on the centre of the screen, 15% lower with respect to Experiment 2, in order to define both contexts symmetrically. For half of the participants the top context was associated with a high proportion of congruent trials (75%) and the bottom context was associated with a low proportion of congruent trials (25%). The reverse was true for the remaining half of the participants.

5.2. Results

5.2.1. Accuracy measures

Data from twelve participants were removed from the analysis because their performance did not reach 60% accuracy in the baseline condition. The remaining 24 subjects (16 women, age range: 18-30 years old, mean age of 20 years old) were included in the complete data analysis.

Mean accuracy for each participant and condition was submitted to a 2 Congruence (congruent, incongruent) by 2 Proportion congruent (high, low) ANOVA, which revealed a significant main effect of Congruence, $F(1,23) = 99.22$, $MSE = 2.04$, $p < .001$, $h^2_p = .81$. Responses on congruent trials (82% correct) were more accurate than responses on incongruent trials (53% correct). The main effect of Proportion congruent was not significant, $F < 1$. In contrast with Experiment 2, the Congruence by Proportion congruent interaction did not reach statistical significance, $F < 1$. The size of the congruence effect was almost identical in the two contexts [28.9% and 29.3% in the low proportion-congruent and high proportion-congruent conditions, respectively, (see Table 1)].

To examine any learning effect of the association between context and likelihood of congruence, data in Experiment 2 were divided into two groups: the first four blocks and the last four blocks. The first ten trials of the first block were again considered as part of practice and were removed from the analysis. A 2 Congruence (congruent, incongruent) by 2 Proportion congruent (high, low) by 2 Block (first half, second half) ANOVA showed that the three-way interaction of Congruence, Proportion congruent and Block did not reach statistical significance $F(1,23) = 1.27$, $MSE = .006$, $p = .27$, $h^2_p = .05$ (see Figure 6).

EXPERIMENT 3

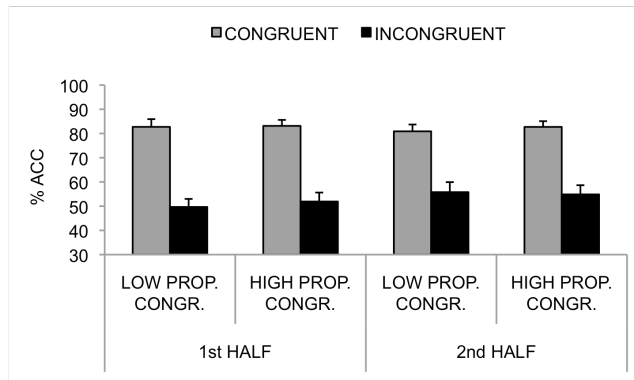


Figure 6. Percentage of correct responses as a function of Congruence, Proportion congruent and Block in Experiment 3. Error bars represent the standard error of the mean across participants.

Neither d' [$t(23)=1.46$, $p=.16$] nor c [$t(23)=1.27$, $p=.22$] indexes differed between high and low proportion-congruent conditions.

5.2.2. Questionnaire measures

Only 4 (17%) of the participants noticed that one of the contexts differed from the other, but no participants were specifically aware of the context manipulation (see Table 2). Instead, the difference between the two contexts was attributed to chance or the fact that one of the locations was simply easier to respond to than the other. Like in Experiment 2, the majority of the participants (79%) considered the information provided by the feedback unreliable (see Table 3).

5.3. Discussion

In Experiment 3, the congruence effect was not modulated by the congruence context manipulation, which contrasts with the results obtained in Experiment 2. One possibility is that the lack of a significant interaction between Congruence and Proportion congruent in Experiment 3 owes to a weak context manipulation, which prevented the association between spatial location and likelihood of congruence. Note that Crump et al. (2006) suggested that a learned association between the context in which the stimuli are presented and the likelihood of congruence is crucial to obtain the context-specific

proportion-congruent effect. To test the hypothesis that the context manipulation was simply too weak to be learned in Experiment 3, the distinction between the two congruence contexts was made more salient in Experiment 4.

6. Experiment 4

6.1. Method

6.1.1. Participants

The participants were 52 undergraduate students (47 women, age range: 18-48 years old, mean age 22 years) who received course credits in exchange for their participation.

6.1.2. Apparatus, stimuli and procedure

These were the same as in Experiment 3, although in this study the white circle (3.01° in diameter) was presented a 6.02° (from the centre of the circle) below or above a central horizontal line, rather than 4.52° below or above a central fixation point.

6.2. Results

6.2.1. Accuracy measures

The data from 16 of the participants² were removed from the analysis because their performance did not reach the 60% accuracy criterion in the baseline condition. The remaining 36 subjects (32 women, age range: 18-48 years old, mean age 22 years) were included in the complete data analysis.

Mean accuracy for the remaining participants in each experimental condition was submitted to a 2 Congruence (congruent, incongruent) by 2 Proportion congruent (high,

² A 2 Congruency (congruent, incongruent) by 2 Proportion congruent (high, low) ANOVA, was conducted with data from all of the participants. In Experiment 2 the interaction between Congruency and Proportion congruent was significant ($F(1,31) = 46.14$, $MSE = .67$, $p < .001$, $h^2_p = .60$). In Experiment 3 the interaction was not significant ($F(1,35) = 2.37$, $MSE = .007$, $p = 1.33$, $h^2_p = .06$). In experiment 4 the interaction was marginally significant ($F(1,51) = 3.70$, $MSE = .01$, $p = .06$, $h^2_p = .06$). Therefore, the results in Experiments 2-4 did not change substantially when including all the participants. In any case, we maintain the exclusion criterion to ensure that our participants understood the task and were responding above chance in the baseline condition.

EXPERIMENT 4

low) ANOVA. This analysis revealed a significant main effect of Congruence, $F(1,35) = 46.23$, $MSE = 1.71$, $p < .001$, $h^2_p = .57$. Responses on congruent trials (77% correct) were more accurate than responses on incongruent trials (56% correct). The main effect of Proportion congruent was significant, $F(1,35) = 5.53$, $MSE = .007$, $p = .024$, $h^2_p = .14$. Participants performed better for the low proportion-congruent context (67,1%) than for the high proportion-congruent context (65,7%). Crucially, the interaction between Congruence and Proportion congruent was significant, $F(1,35) = 4.75$, $MSE = .01$, $p = .036$, $h^2_p = .12$. Participants performed more accurately on congruent trials than on incongruent trials for both Proportion congruent conditions, $F(1,35) = 33.42$, $MSE = .71$, $p < .001$ and $F(1,35) = 54.39$, $MSE = 1.01$, $p < .001$, high and low respectively. However, in line with the results in Experiment 2, the size of the Congruence effect was smaller for the low proportion-congruent context (20%) than for the high proportion-congruent context (24%), $t(36) = 2.18$, $p < 0.04$, (see Table 1). A priori comparisons revealed that this difference was due to an increased accuracy on incongruent trials (3,4%), $F(1,35) = 7.89$, $MSE = .02$, $p = .008$, in the low proportion-congruent context relative to the high proportion-congruent context. There was no statistical difference on congruent trials between the two contexts, $F < 1$.

A 2 Congruence (congruent, incongruent) by 2 Proportion congruent (high, low) by 2 Block (first half, second half) ANOVA was also conducted in Experiment 4 (see Figure 7). The first ten trials were excluded from the analysis. The three-way interaction between Congruence, Proportion congruent and Block did not reach statistical significance, $F(1,35) = 1.79$, $MSE = .011$, $p = .19$, $h^2_p = .048$. However, posterior analyses showed a significant interaction between Congruence and Proportion congruent for the second half of the experiment, $F(1,35) = 5.16$, $MSE = .036$, $p < .03$, $h^2_p = .13$, while this interaction was not significant for the first half of the experiment $F < 1$. A priori comparisons revealed that, in line with results in Experiment 2, a significant difference between accuracy on incongruent trials of the second half of the experiment $F(1,35) = 12.02$, $MSE = .052$, $p < .002$, with a better performance (5.36%) in the low proportion-congruent context than in the high

proportion-congruent context. In the first half of the experiment this difference was not significant, $F < 1$.

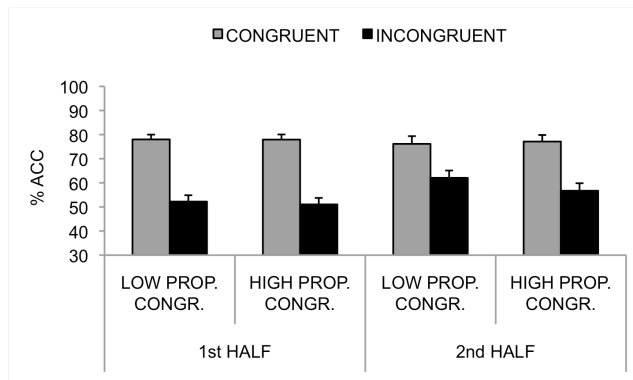


Figure 7. Percentage of correct responses as a function of Congruence, Proportion congruent and Block in Experiment 4. Error bars represent the standard error of the mean across participants.

In contrast to the results in Experiment 2, pair-wise two-tailed t-tests conducted between high and low proportion-congruent for d' and c indexes showed that d' was significantly higher on the low proportion-congruent (2,24) condition than in the high proportion-congruent condition (1,96), $t(35) = 2.35$, $p = .02$, while no differences were obtained in c (0.28 for high, 0.32 for low, $t < 1$).

6.2.2. Questionnaire measures

Only 6 participants (17%) noticed a difference between the two contexts, although none of them was aware of the specific proportion-congruent manipulation (see Table 2). Instead, the perceived difference was explained with comments related to tiredness and body position.

To examine differences on performance between participants who reported any difference between the two context and participants who did not, a 2 Congruence (congruent, incongruent) by 2 Proportion congruent (high, low) ANOVA was conducted for each group separately, given that the large disparity in the number of participants who belong to each group (6 difference and 30 no difference; see Figure 8). In line with Experiment 2, the analyses revealed a significant interaction between Congruence and

EXPERIMENT 4

Proportion congruent only for the group who did not report any difference between the two contexts, $F(1,29) = 5.87$, $MSE = .018$, $p < .03$, $h^2_p = .168$, while the interaction did not reach statistical significant for the group of participants who noticed differences between both context, $F < 1$. In line with Experiment 2, a priori comparisons revealed a statistical significant difference between performance on incongruent trials in the low proportion-congruent context (59%) with respect to the high proportion-congruent context (55%), for participants who did not report any difference between the two contexts $F(1,29) = 11.70$, $MSE = .027$, $p < .002$. Performance between incongruent trials from the two contexts did not reach statistical significant in the group of participants who reported differences between the two contexts, $F < 1$.

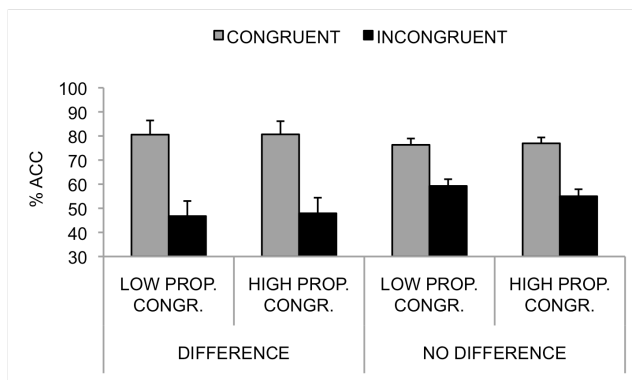


Figure 8. Percentage of correct responses as a function of Congruence, Proportion congruent and Context difference in Experiment 4. Error bars represent the standard error of the mean across participants.

Once again, participants' reports at the end of the experiment revealed that feedback information was considered unreliable by the 67% of the participants (see Table 3).

6.3. Discussion

Experiment 4 revealed that increasing the spatial separation between the two congruence contexts relative to Experiment 3 produced the expected result: a larger congruence effect in the high proportion-congruent context than in the low proportion-

congruence context. Like in Experiment 2, the data from the questionnaire suggests that the effect was not due to voluntary conscious control. Furthermore, given that observers could not predict whether the oncoming trial belonged to the high or low proportion-congruent context, they could not prepare themselves in advance to exert more or less attentional control. We argue that the particular location where the stimuli were presented became associated with a particular level of congruence, which in turn led to different control set being cued in the two contexts.

In Experiment 4, a significant increment in sensitivity (d') in the low proportion-congruent context was obtained. This result suggest that the congruence manipulation modulated audiovisual interactions at a perceptual level of information processing, when different proportions of congruent and incongruent trials were presented mixed up in the same block of trials. This perceptual sensitivity enhancement could be related to the better performance of incongruent trials on the low proportion-congruent condition. However, given that the same modulation of the d' index was not obtained in Experiment 2, this result should be taken with caution. A potential, and speculative, explanation refers to the trial by trial shifts in the attentional control set that would be required in Experiment 4 in contrast to Experiment 2, in which a constant proportion of congruence was presented on each block of trials. This could explain the differential influence in the d' index. Importantly, the criterion index (c) in Experiment 4 was not modulated by the context of congruence manipulation.

Feedback in Experiments 1-4 was presented on the assumption that it would give information about the proportion of congruence, facilitating the association between context and likelihood of congruence. However, it is clear that participants were not aware of the proportion of congruence manipulation in any of the experiments reported in this manuscript. What's more, once again, the majority of participants in Experiment 4 considered the feedback information as unreliable. Therefore, feedback does not seem to have played a key role in our findings.

7. General Discussion

The purpose of this study was to examine the effect of congruence context on the interaction of auditory and visual stimuli. The audiovisual interaction was measured by means of a task in which participants had to discriminate the duration of a visual stimulus while ignoring the duration (congruent or incongruent) of a synchronously-presented auditory stimulus. The congruence context was manipulated by varying the relative proportion of congruent and incongruent trials, between blocks in Experiment 2, and within blocks in Experiments 3 and 4.

The results of Experiment 1 showed that our stimuli were suitable to investigate crossmodal congruence effects, similar to those reported in previous studies (e.g., Andersen, Tippana & Sams, 2004; McGurk & MacDonald, 1976). Note that the results of Experiment 1 are consistent with the outcome of a recent study reported by Klink et al. (2011) who showed that the perceived duration of a visual stimulus depended on the duration of an auditory stimulus presented synchronously. Accordingly, given that audition dominates vision in the processing of temporal information, we would argue that the perceived visual duration in our study was biased towards the perceived duration of the auditory input (e.g., Romei, De Haas, Mok & Driver, 2011; Shams, Kamitani & Shimojo, 2000, 2002; Walker & Scott, 1981). SDT analyses pointed that the congruence effect measured here reflects crossmodal interactions occurring at perceptual and post-perceptual level of stimulus processing, in accordance with previous accounts (e.g., Sanabria et al., 2007)

The results of Experiment 2 revealed that the proportion of congruent trials, manipulated in separate blocks of trials, influenced the size of the congruence effect. Crucially, Experiment 4 demonstrated that the proportion of congruence could influence the congruence effect even when different proportions were associated with two contexts presented within the same block of trials. Taken together, these results suggest that audiovisual interactions measured by the congruence effect are prone to top-down

attentional modulations related to the context of congruence, in contrast to previous accounts (see Shore and Simic, 2006).

The results reported here are in accord with those of Crump and colleagues (2006); that is, a larger crossmodal congruence effect was observed for the context in which most of the trials were congruent. We suggest that, in terms of cognitive control, this effect hinges on context-sensitive adjustments as a function of audiovisual congruence. In our study, for each location (either above or below the fixation point), the particular congruence proportion defined whether the auditory input shared its duration with the visual target input. As such, we would argue that the low proportion-congruent context triggered an attentional set for filtering the visual input from the temporally incongruent auditory input. In contrast, in the high proportion-congruent context the auditory input shared the same temporal parameters than the visual input in most of the occasions, and so that attentional filtering would not be required, or, at least, not at the same extent as in the low proportion-congruent condition. It appears then that context can drive the attentional set that modulates the way in which audiovisual inputs interact. The outcome of Experiment 4 demonstrated that these shifts in the attentional set were highly flexible and could occur on a trial by trial basis.

The results of Experiments 3 and 4 suggest that distinct location contexts are necessary to learn the association between those contexts and different congruence proportions. In Experiment 3, it seems likely that the two locations were not sufficiently distinct to be treated as separate contexts, and instead were processed as one. The greater distinctiveness of the two contexts in Experiment 4 increased the likelihood that they would be selectively attended to as separate contexts. Thus, participants were able to learn the association between location and proportion congruence, which resulted in the context-specific proportion-congruent effect observed in Experiment 4 but not in Experiment 3.

GENERAL DISCUSSION

The results in Experiments 3-4 are consistent with the idea that implicit learning can depend on attention to the task relevant dimensions (e.g., Jiménez & Méndez, 1999). Jiménez and Méndez showed the difficulty to associate a shape with a location if the shape was task-irrelevant, given that attention was not focused on shape dimension. Crucially, the distinct location contexts used in Experiment 4 were more likely to require shifts of attention than those used in Experiment 3, and thus the association between location and proportion congruent was also more likely to be learned in Experiment 4 than in Experiment 3 (see also Crump et al., 2008). The improvement of participants' performance on incongruent trials in the low proportion-congruent condition (with respect to the high proportion-congruent condition) in the first half of the experiment as compared to the second half in Experiments 2 and 4, strengthens the argument that the association between context and proportion of congruence has to be learned. The fact that this improvement was absent in Experiment 3 bolsters the idea that attentional shifts are necessary to set this association, not only experience within a particular context of congruence. Given that participants were not aware of the association between the location of the stimuli and audiovisual congruence, we would argue that this association occurred implicitly.

The subjective measures collected in our study suggest that the attentional mechanism responsible for the reduction of the congruence effect in the low proportion of congruence context was involuntarily triggered by the stimulus characteristics, which entailed rapid shifts after stimulus onset. In fact, the large majority of participants in Experiments 2 and 4 did not notice the context manipulation. Moreover, when the group of participants was split into those who noticed differences between the two contexts of congruence and those who did not, the former showed a decreased modulation of the congruence. Noticing a difference between the two contexts (whatever that difference was) might have resulted in participants using different task strategies for each context of

congruence. This would have interfered with the control mechanism automatically triggered by the onset of the stimulus at each particular context.

Previous studies have shown that contextual cues can control selective attention processes during online performance in a fast and stimulus-driven manner (see Egner, 2008). In line with this idea, each location in Experiments 2 and 4 may constitute a contextual cue, which triggers the attentional control needed to select relevant information. Given that the location contexts were mixed at random across trials in Experiment 4, this form of control would not be implemented prior to stimulus onset, but instead selection of a particular set of control would be cued by the onset of the stimuli at either one context or the other. By this view, context can cue control shifts involuntarily during on-line performance.

In sum, our study shows, for the first time, that the context of congruence, resulting in shifts in attentional control set, can modulate the outcome of audiovisual interactions. We propose that this modulation was caused by selective attention processes involved in the association between context and likelihood of congruence, presumably in an “automatic” context-driven manner. We therefore contribute to the ongoing debate regarding the role of attention on multisensory perception suggesting a crucial role of attention on crossmodal processing.

This study was supported by a Discovery research grant from the Natural Sciences and Engineering Research Council of Canada to David I. Shore, a Discovery research grant from the Natural Sciences and Engineering Research Council of Canada to Bruce Milliken, a student scholarship (AP2008-03662) from the Ministerio de Educación y Ciencia to Beatriz R. Sarmiento, grants SEJ2007-63645 and PSI2010-19655 from the Ministerio de Educación y Ciencia and SEJ-6414 from the Junta de Andalucía to Daniel Sanabria. We thank Sarah Lade for her help with data collection.

**CAPÍTULO 4. SERIE EXPERIMENTAL II.
ELECTROPHYSIOLOGICAL EVIDENCE OF
AUDIOVISUAL INTERACTIONS UNDER
DIFFERENT CONTEXTS OF STIMULUS
CONGRUENCE**

El contenido de este capítulo está en preparación como Sarmiento, B.R., Murray, M.M. and Sanabria, D. Electrophysiological Evidence of audiovisual interactions under different contexts of congruence.

1. Abstract

This study analyzed high-density event-related potentials within an electrical neuroimaging framework to provide insights regarding the interplay between attention and multisensory processes. Specifically, we identified the spatio-temporal brain mechanisms by which the probability of temporally congruent and task-irrelevant auditory information influences stimulus processing during a visual duration discrimination task. The spatial position (top/bottom) of the visual stimulus was indicative of the probability of the co-occurring auditory stimulus being congruent vs. incongruent in its duration. Electrophysiological analyses showed a modulation of the audiovisual processing as a function of the proportion of congruence - i.e., greater strength for congruent than incongruent inputs in the location where most of the trials were congruent - over the ~163-227 ms post-stimulus period. Differential neural activity in primary areas (e.g., auditory cortex) as well as high-order areas (e.g., left middle frontal gyrus) was observed as a function of the context of congruence (i.e., spatial position) on that interval. At the ~252-272 ms interval audiovisual interactions depended both on the proportion of congruence of the current trials and that of the preceding trial in line with the behavioural results. Importantly, a control oddball task revealed greater strength for responses to audiovisual stimuli than to the summed unisensory inputs [$AV > (A+V)$]. This result was accompanied by an activation of multisensory areas (i.e., left cuneus and middle occipital gyrus) for the audiovisual stimuli. Notably, the oddball task also showed that congruent and incongruent stimuli were equally processed when congruence was not a relevant dimension. In sum, this study demonstrated that top-down processes influence audiovisual integration over different stages of processing. Crucially, such top-down processes are triggered in response to the characteristics of the context (e.g., probability of congruence), and only related to relevant dimensions (e.g., incongruity) for the task at a hand.

2. Introduction

The interplay between attention and multisensory processing (e.g., Spence & Driver, 2004) has remained controversial over the last two decades. Several studies have suggested that multisensory integration occurs at a pre-attentive stage and that it is largely impervious to attentional influences (e.g., Driver, 1996; Bertelson & de Gelder, 2004). This notion has been challenged by other behavioural (e.g., Sanabria, Soto-Faraco & Spence, 2007a), event-related potential (ERP; e.g., Talsma, Doty & Woldorff, 2007), and neuroimaging (e.g., Fairhall & Macaluso, 2009) findings supporting the claim that attention can affect multisensory outcomes (see van Atteveldt, Murray, Thut & Schroeder, 2014 for a review).

Recently, Talsma, Senkowski, Soto-Faraco & Woldorff (2010) proposed a bidirectional framework to account for the relationship between attention and multisensory processing. According to this model, salient multisensory stimuli tend to capture attentional resources in a bottom-up fashion. But, whenever incongruity between stimulus representations from two sensory modalities is detected, top-down control processes are recruited to prevent the integration of multisensory inputs. For instance, Talsma & Woldorff (2005) presented evidence in favour of an early modulation of ERPs associated with the presentation of multisensory inputs, amplitudes were enhanced when audiovisual stimuli were selectively attended. Therefore, it was suggested that top-down control was needed to select appropriate (to-be-integrated) information in situations where competing sensory representations were activated.

The role of top-down control on stimulus processing has been broadly studied in cognitive control research using conflict paradigms as Flanker (e.g., Corballis & Gratton, 2003; Appelbaum, Carsten, Boehler, Chen & Woldorff, 2011) or Stroop (e.g., Egnér & Hirsch, 2005; Crump, Gong & Milliken, 2006) where congruence between two different dimensions of unisensory stimuli is manipulated.

Recently, Sarmiento, Shore, Milliken & Sanabria (2012) showed that different stimulus contexts (i.e., locations), wherein the likelihood of congruence between the duration of visual and auditory stimuli was manipulated (hereafter context of congruence), influenced the outcome of audiovisual interactions in a trial-by-trial manner. Specifically, participants discriminated the duration (short or long) of a visual stimulus (a white circle), which could be accompanied by a congruent or incongruent auditory distractor (a white noise burst). Target and distractor matched in duration most of the occasions in the location associated with a high proportion of congruent inputs, whereas they were more often mismatched in the low proportion-congruent location.

Sarmiento et al. (2012) used signal detection theory to assess whether top-down control influenced audiovisual interaction at a perceptual (viz. sensitivity; d'), post-perceptual (viz. criterion; c) or both stages of stimulus processing. Although auditory and visual inputs interacted at both levels of processing (see Sanabria, Spence & Soto-Faraco, 2007b, for related results), attentional modulation seemed to occur at a perceptual stage, showing an enhanced sensitivity for visual stimuli presented in the low proportion-congruent context. The authors argued that the target onset location triggered attentional adjustments responsible for selecting the suitable incoming information, resulting in larger crossmodal interaction, measured as the difference in performance for congruent and incongruent audiovisual stimuli, at the location with high likelihood of congruent stimuli than at the location where most of the trials were incongruent. This result suggested that top-down control was applied to a larger extent at the location with a greater number of incongruent trials than at the high congruence context, in order to perform the task accurately by avoiding the auditory influence on visual target processing. To our knowledge, this result showed for the first time evidence for the influence of on-line top-down control shifts in crossmodal outcomes in a stimulus congruence paradigm.

In line with Talsma et al.'s (2010) framework, Sarmiento et al. (2012) findings suggest that in multisensory competing situations attentional mechanisms are activated prior to

INTRODUCTION

multisensory interactions in order to obtain an efficient processing of sensory inputs. In addition, it appeared that the relative context of congruence associated with a particular location was responsible for activating the appropriate control settings. Findings from several studies using unisensory congruence tasks support the idea of this on-the-fly control adjustment approach that can be triggered by contextual informative cues (e.g., Crump et al., 2009) or by mere context repetition (e.g., King et al., 2012, Spape and Hommel, 2008).

There are two main important issues that emerged from Sarmiento et al.'s (2012) study and which are the focus of the present study. First, we determined the way auditory and visual inputs interact, irrespective of the context of congruence or the particular tasks demands. This was assayed here by contrasting EEG responses to multisensory stimuli (i.e., AV) with summed responses to unisensory conditions (i.e., A+V). Following this approach, non-linear neural response interactions have been documented within the initial 200 ms post-stimulus onset (e.g., Giard and Peronnet 1999; Molholm et al., 2002, Fort et al., 2002; Cappe et al., 2010, 2012). The second critical question concerns the stage of stimulus processing at which top-down control influenced auditory-visual interactions. Several studies have demonstrated that multisensory processing can be modulated by attention across various stages of stimulus processing. As mentioned above, electrophysiological evidence has shown that attention to the location of auditory-visual stimuli influences ERPs below 200 ms after target onset as well as at later stages (~200-500 ms) (see Talsma & Woldorff, 2005). Other EEG methods have pointed in the same direction. For instance, Senkowski, Talsma, Herrmann & Woldorff (2005), using a time-frequency analysis approach found spatial attention effects at around 50 ms after the onset of the auditory-visual stimulus.

Here, we conducted a high-density ERP study and we applied multivariate signal analysis techniques for EEG referred to as electrical neuroimaging that distinguish between modulations in response strength, topography and latency (e.g., Murray et al., 100

2008; Michel and Murray, 2012). In order to assess multisensory integration we used an oddball task wherein visual and auditory unisensory stimuli as well as audiovisual multisensory stimuli were presented. We expected to observe differences in EEG responses to multisensory stimuli in contrast with summed responses to unisensory inputs. In addition, the same paradigm as in Sarmiento et al. (2012) was used to determine the time course of the attentional modulation over audiovisual interactions as a function of the context of stimulus congruence. At a behavioural level, we expected to replicate our previous results. At an electrophysiological level, we expected to observe a multisensory facilitation effect as well as an early modulation of the processing of auditory-visual stimuli as a function of the context of congruence.

3. Method

3.1. Participants

The participants were 28 undergraduate students from the University of Granada (19 females; age range: 18-28 years old; mean age of 24 years) who received course credits or cash payment of 20 Euro in exchange of their participation. All of the participants reported normal hearing and normal or corrected-to-normal vision and gave their informed consent to participate in the study conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

3.2. Apparatus and stimuli

The experiment was conducted on an Intel Core 2 Duo PC with a 17-in. LCD monitor. E-Prime software was used for stimulus presentation and response collection (Psychology Software Tools; www.pstnet.com).

The visual stimuli consisted of a white circle (3.01° in diameter), a white equilateral triangle (3.01° in height) and a white central cross that served as the fixation point. The

METHOD

horizontal line of the cross (11.52° long) split the screen into upper and lower halves. Two loudspeakers, positioned on each side of the computer screen, were used to present the auditory stimuli, which consisted of a white noise burst and a 600 Hz pure tone (60 dB[A] measured at ear level; see Figure 9).

3.3. Procedure

Participants sat in a comfortable chair at approximately 57 cm from the computer monitor in a silent, dark and electrically shielded room. The instructions that explained the task were displayed on the computer screen. The experiment began with an unisensory discrimination task in which participants were asked to discriminate the duration (short or long) of a white circle displayed above or below (6.52° from the centre of the circle) the fixation point (i.e., the white horizontal line) for either 100 ms or 180 ms. The unisensory block (16 short and 16 long unisensory trials) was used to ensure that participants understood the task and were responding above 60% of accuracy overall.

After the unisensory task participants put on an elastic electrode-cap for the EEG recording (see EEG acquisition section below).

Next, participants performed an oddball task in which either a white triangle or a 600 Hz pitch had to be detected via button-press ("3" key of a response box). The standard stimuli, which did not require a response, were a white circle and a white noise burst. The visual stimuli could appear above or below the fixation cross (6.52° from the centre of the circle). Participants performed two blocks of 216 trials each, 11% of which were oddballs (i.e., 24 trials). Oddballs and standards could last for 100 or 180 ms and could be presented alone (unisensory stimuli) or synchronously with a stimulus of the other sensory modality (multisensory stimuli). The visual and auditory oddballs were never presented together. This manipulation gave rise to unisensory trials (the visual or auditory stimulus was presented alone), multisensory congruent trials (the visual stimulus was of the same duration as the auditory stimulus) and multisensory incongruent trials (the visual stimulus

was of a different duration than the auditory stimulus). Unisensory and multisensory stimuli, short and long stimuli, as well as congruent and incongruent stimuli had the same probability of appearance. Participants had to detect the oddballs when they appeared alone as well and when they were presented accompanied by a standard stimulus from the other modality. Feedback was only displayed whenever participants made an error. The oddball task was used as a baseline condition to assess multisensory integration in congruent and incongruent audiovisual stimuli when no response was required (i.e., standard stimuli).

The oddball task was followed by the visual duration discrimination task in which participants were asked to discriminate the duration of a white circle displayed below or above the fixation point for either 100 ms or 180 ms, while ignoring the synchronous presentation of a white noise burst that could last for 100 ms or 180 ms. No unisensory trials were presented in the visual duration discrimination task. As in the case of the oddball task, congruent and incongruent trials emerged from the four possible stimuli combinations. The crucial manipulation consisted of the inclusion of two contexts of congruence defined by the two halves of the computer screen (upper/lower). For half of the participants the top context was associated with a high proportion of congruent trials (80%) and the bottom context was associated with a low proportion of congruent trials (20%). The reverse was true for the remaining half of the participants. Participants completed 8 multisensory practice trials and 8 blocks of 120 multisensory trials each, distributed as follows: 48 congruent trials and 12 incongruent trials for the high proportion-congruent context; 12 congruent and 48 incongruent trials for the low proportion-congruent context. The congruent and incongruent stimulus had the same probability of appearance. This was also true for short and long trials. Accuracy, rather than response speed, was stressed; so participants had no response time pressure. Responses were collected using a serial response box. Half of the participants were told to press the “1” key if the circle was short

METHOD

in duration and the “4” key if the circle was long in duration. The reverse stimulus-response mapping was used for the remaining participants.

At the beginning of each trial, participants were presented with the fixation cross displayed in white against a black background for a random duration between 500 and 1000 ms. This fixation cross remained on and steady for the whole trial. Participants were encouraged to fixate and to avoid eye movements and blinks during the presentation of the stimuli. Then, the visual and auditory stimuli were presented for either 100 ms or 180 ms, and could be congruent or incongruent in duration. Feedback regarding response accuracy was provided to the participants for 500 ms. The next trial began between 1500 and 2000 ms after the feedback. The sequence of events is schematized in Figure 9.

At the end of the experiment, participants completed a questionnaire to evaluate whether or not they were aware of the congruence manipulation. They were asked whether they perceived any difference between the two onset locations during the visual duration discrimination task, and what that difference might be.

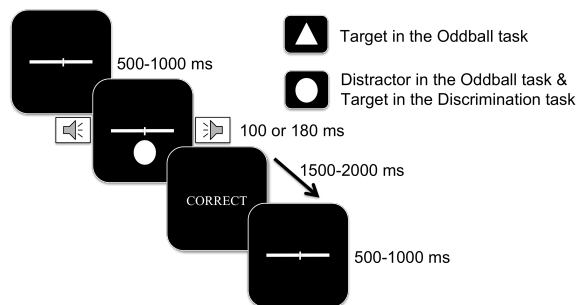


Figure 9. Schematic illustration of the set up used in the Experiment.

3.4. EEG acquisition

Continuous EEG was acquired at 1000 Hz through a Geodesic Sensor Net of 129 Ag/AgCl electrodes referenced to the vertex channel (Tucker, Liotti, Potts, Russell, & Posner, 1994) and connected to an AC-coupled high-input impedance amplifier (200 M Ω , Net Amps™, Electrical Geodesics, Eugene, Oregon). Impedances were kept below 50k Ω , as recommended for these amplifiers. The signal was acquired with a 0.01-100Hz elliptical

band-pass filter. Gain and zero calibration were performed prior to the start of every recording. The head-coverage included sensors lateral to and below both eyes to monitor horizontal and vertical eye movements (electrooculogram, EOG).

3.5. ERP analyses

The EEG was 0.1-30Hz band-pass filtered offline. Epochs of EEG were segmented from 200 ms pre-stimulus to 800 ms post-stimulus onset. The segmented epochs were submitted to automated software processing for identification of artifacts. Epochs containing eye movements or blinks (± 70 μV relative to baseline in EOG channels) were rejected. Individual channels containing other sources of transient noise (± 80 μV relative to baseline in any channel) were replaced using a trial-by-trial basis with a spherical interpolation algorithm (Perrin, Pernier, Bertrand, & Echallier, 1989). The epoch was discarded when more than 10 channels were deemed artifact-contaminated. Additionally, the data were visually inspected to ensure that all artifacts had been detected. A minimum of 34 trials per subject and condition were required to ensure an acceptable signal-to-noise ratio. Epochs of EEG were averaged for each stimulus condition and from each subject to calculate ERPs according to each of the conditions in the different tasks. Data from ERPs were baseline corrected using the period of 200 ms preceding stimulus onset. ERPs were recalculated against the average reference. All analyses were conducted using the freeware Cartool (<http://sites.google.com/site/fbmlab/cartool>; Brunet et al., 2011) and the STEN toolbox (<http://www.unil.ch/fenl/page99386.html>). The analysis strategy we used followed a multi-step analysis procedure referred to as electrical neuroimaging (e.g., Murray et al., 2008).

Electrical neuroimaging allowed us to identify multisensory effects using both local and global measures of the electric field at the scalp. This procedure distinguishes between effects following from modulations in the strength of responses of statistically indistinguishable brain generators and alterations in the configuration of these generators

METHOD

(viz. the topography of the electric field at the scalp), as well as latency shifts in brain processes across experimental conditions (e.g., Michel et al., 2004, 2009; Murray et al., 2008; Michel & Murray, 2012; Altieri et al., 2013; Stevenson et al., 2014). Additionally, we used the local autoregressive average distributed linear inverse solution (LAURA; e.g., Grave de Peralta Menendez et al., 2001, 2004) to visualize and statistically contrast the likely underlying sources of effects identified in the preceding analysis steps.

Global modulations in the strength of the electric field at the scalp were quantified by calculating the global field power (GFP; e.g., Lehmann and Skrandies, 1980) for each subject and stimulus condition. This measure represents the spatial standard deviation of the electric field at the scalp at a given time point, and constitutes a reference-independent measure of the ERP amplitude (e.g., Murray et al., 2008; Koenig and Melie-Garcia, 2010). Stronger electric fields result in larger GFP values. GFP values were baseline corrected in order to legitimately compare between conditions with different numbers of trials. GFP modulations were analyzed using a millisecond-by-millisecond ANOVA and when warranted paired t-tests. Both analyses were performed in conjunction with a 12 contiguous data-point temporal criterion for significant effects to correct for temporal auto-correlation (only those effects where p-values met the statistical threshold ($p \leq 0.05$) for more than 12 contiguous time points (~12 ms at 1000 Hz sampling) were considered reliable; inspired by Guthrie & Buchwald, 1991).

In order to statistically determine periods of topographic modulations independently of ERP strength, we calculated the global dissimilarity (DISS; e.g., Lehmann and Skrandies, 1980) as a function of Context and Congruency. This parameter is an index of configuration differences between two electric fields, independent of their strength, whose values range from 0 to 2, where 0 indicates topographic homogeneity and 2 indicates topographic inversion (e.g., Murray et al, 2008).

In order to identify stable periods of electric field topography (hereafter template maps), the collective post-stimulus group-average ERPs were subjected to hierarchical

clustering. The optimal number of stable ERP clusters (i.e., the minimal number of maps that accounts for the greatest variance of the dataset) was determined using a modified Krzanowski-Lai criterion (e.g., Murray et al., 2008). The clustering makes no assumption on the orthogonality of the derived topographic template maps (e.g., De Lucia et al., 2010a, 2010b; Pourtois et al., 2008). Template maps identified in the group-level ERP were entered into a fitting procedure wherein each template map was compared with the moment-by-moment scalp topography on the individual subjects' ERPs from each condition and was labeled according to the one with which it best correlated spatially. This fitting procedure allows determining the total amount of time a given template map was observed for a given condition across subjects. Statistical analysis of these values was performed with an ANOVA. The results of this analysis revealed whether a given ERP was more often described by one template map vs. another, which implied different stable map configurations across conditions at the same moments in time.

At a neurophysiologic level, the topographic procedure detailed above revealed when the experimental conditions activated distinct sets of brain networks, because electric field changes are indicative of changes in the underlying generator configuration (e.g., Lehmann, 1987). Additionally, this method was insensitive to pure amplitude modulations across conditions because the data were first normalized by their instantaneous GFP. As above, only effects where p values were <0.05 for at least 12 contiguous time points (~12 ms) were considered reliable.

3.6. Source Estimations

The intracranial generators underlying the surface electric field were estimated using a distributed linear inverse solution based on a Local Auto-Regressive Average (LAURA) model comprising biophysical laws as constraints (e.g., Grave de Peralta Menendez et al., 2001, 2004). LAURA uses a realistic head model with a solution space of 4024 nodes, selected from a 6 x 6 x 6 mm grid equally distributed within the gray matter of the Montreal

RESULTS: ODDBALL TASK

Neurological Institute's (MNI's) average brain. This algorithm selects the source configuration that better mimics the biophysical behaviour of electric vector fields, that is, the estimated activity at one point depends on the activity at neighboring points according to electromagnetic laws (e.g., Grave de Peralta & Gonzales Andino, 2002). Prior to calculation of the inverse solution, the ERP data were down-sampled and affine-transformed to a common 111-channel montage. The time period for which intracranial sources were estimated and statistically compared between conditions was defined by the results of the abovementioned topographic analysis. Statistical analyses of source estimations were performed by first averaging the ERP data across time to generate a single data point for each participant and condition. Paired t-test or an ANOVA were performed at each of the 4024 source nodes in the inverse solution space (using across-subjects variance). Only nodes exceeding the spatial extent criterion of at least 17 contiguous significant nodes ($p \leq 0.0002$) were considered significant (see Thelen et al, 2012 for similar spatial criterion). This spatial criterion was determined using the AlphaSim program (<http://afni.nimh.nih.gov>) and assuming a spatial smoothing of 6mm full-width half maximum. The results of the analyses were rendered on the Montreal Neurologic Institutes's average brain with the Talairach and Tournoux (1988) coordinates of the largest statistical differences within a cluster indicated.

4. Results

The data from two participants were removed due to technical issues with Net Station Software. Additionally, six participants did not reach a minimum of 34 trials per condition due to excessive EEG artifacts, and the data were also removed. The data from the remaining 20 participants (12 women; age range: 18-28 years old, mean age 23) were included in the complete data analysis.

4.1. Oddball Task

4.1.1. Behavioural data

Overall accuracy across participants on the oddball task was 99.77%, with a mean reaction time of 340 ms. This high performance assured that participants were actually engaged in the oddball task.

4.1.2. Surface ERP data

Oddball task analyses focused on determining differences in the neural activation between the ERP in response to the multisensory condition and the summed ERPs in response to the unisensory auditory and visual conditions (denoted by AV and A+V, respectively). Only trials for standard stimuli were included in the analyses since they constituted the pool of stimuli used for the multisensory discrimination task, which is the main manipulation of this study.

The GFP from the AV and A+V responses were compared statistically by paired t-tests as a function of time. Results included reliable strength differences over two periods of time (i.e., from ~114 to 154 ms and from ~177 to 209 ms; see Figure 10). At both intervals, a stronger response to multisensory than to summed unisensory stimuli was observed. It is noteworthy that no differences between congruent and incongruent stimuli were observed.

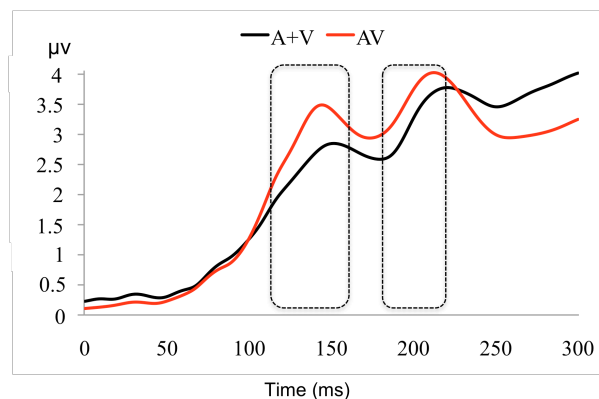


Figure 10. Group-average GFP waveforms and results of t-tests on the GFP as a function of time ($p < 0.05$; > 12 ms). Dotted boxes represent intervals of statistical significant differences.

RESULTS: ODDBALL TASK

Next, we tested for topographic differences between the two conditions.

A hierarchical spatio-temporal cluster analysis was conducted to determine the most dominant template maps appearing in the group-averaged ERPs across time and in both AV and A+V conditions. This procedure identified 22 template maps explaining 97.93% of the variance of the concatenated group-averaged ERP data set. Several maps were identified that appeared to differentially predominate in response to the AV vs. A+V conditions over the ~115-154 ms post-stimulus period. This pattern observed in the group-averaged ERPs was statistically assessed in the single-subject data using a spatial correlation fitting procedure. The amount of time each template map yielded a higher spatial correlation with the single-subject data from each condition was quantified over both intervals and submitted to a repeated-measures ANOVA using within-subject factors of Stimulus condition (AV, A+V) and Template map. A significant interaction between Stimulus condition and Template map was observed over the ~115-154 ms post-stimulus interval, $F(1,19) = 10.19$, $MSE = 405$, $p < 0.005$, $\eta_p^2 = .35$. Further comparisons showed that over this period one map (See Figure 11a, green map) predominated the responses to AV condition $F(1,19) = 10.19$, $MSE = 202.5$, $p < 0.005$, while the other map accounted for the responses to the A+V condition (See Figure 11a, blue map), $F(1,19) = 10.19$, $MSE = 202.5$, $p < 0.005$ (See Figure 11b).

The hierarchical spatio-temporal cluster analysis was also conducted in the group-averaged ERPs across time and both congruent and incongruent conditions. No different template maps were identified between the conditions within the ~115-154 ms post-stimulus period.

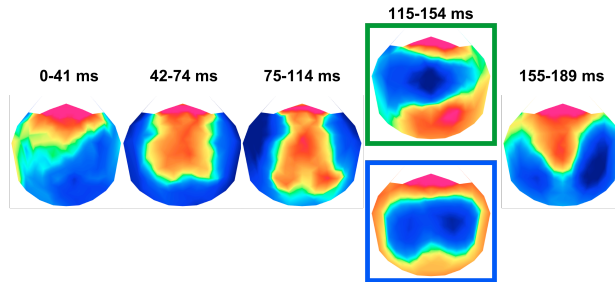
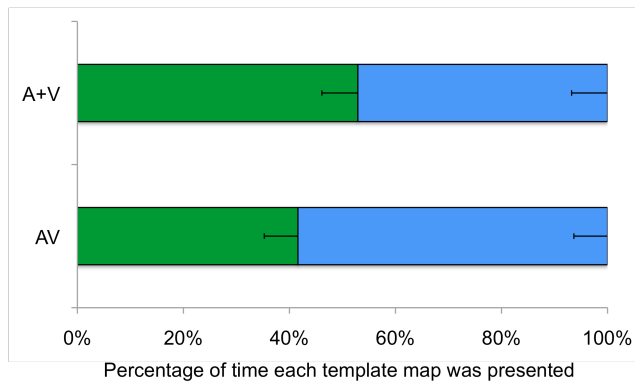
a**b**

Figure 11. a Template maps from topographic cluster analysis. **b.** Results of the spatial correlation fitting over 115-154 ms period in the single-subject data from each condition. Error bars represent the standard error of the mean.

4.1.3. Source estimations

The source estimation from the 115-154 ms time period was statistically analyzed (see Method section) to estimate the intracranial generators contributing to the differences between the AV and A+V conditions.

Results revealed a reliable difference between both conditions in two main clusters within the occipital lobe (See Figure 12). The clusters that showed greater activity for the AV condition than for the A+V condition were located in the left cuneus (Max: -11, -85, 28 mm; BA 19) and in middle occipital gyrus (Max: -35, -87, 6 mm; BA 19).

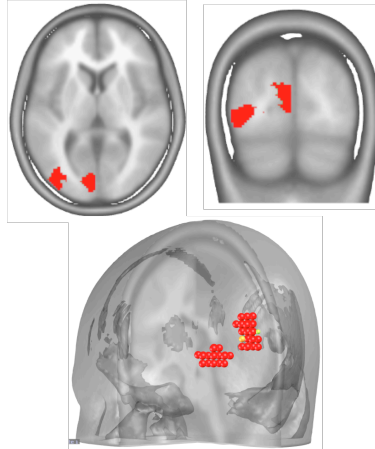


Figure 12. Results of the source estimations from the 115-154 ms post-stimulus onset time window. Note that the positive values (warmer colors) indicate that the clusters are more active in the AV condition than in the A+V condition.

The aim of the oddball task was to assess whether the audiovisual stimuli used in our task exhibited non-linear neural response interactions. Superadditive strength modulations were observed at ~114-154 and ~177-209 ms, and topographic modulations were evident over the ~115-154 ms post-stimulus period. Brain areas involved in these interactions were identified with a distributed source estimation method and yielded in multisensory integration areas (e.g., Calvert et al., 2000; Dhamala et al., 2007; Cappe et al., 2012; Diaconescu et al., 2013; Pillai et al., 2013).

4.2. Visual Temporal Discrimination Task

4.2.1. Behavioural data

Mean response accuracy for each participant and condition were submitted to a 2 Context of congruence (high congruent, low congruent) x 2 Congruence (congruent, incongruent) ANOVA, which revealed a significant main effect of Congruence, $F(1,19) = 42.60$, $MSE = .83$, $p < .001$, $\eta_p^2 = .69$. Responses on congruent trials (83% correct) were more accurate than responses on incongruent trials (62% correct). Both the main effect of

Context of congruence and the interaction between Context of congruence and Congruence failed to reach statistical significance, both $F_s < 1$.

To investigate whether context specific proportion-congruent effect relied on trial-by-trial shifts in control setting we performed a 2 Context transition (repetition, switch) x 2 Context of congruence (high congruent, low congruent) x 2 Congruence (congruent, incongruent) ANOVA (see King et al., 2012, for a similar approach). The first trial of each block was excluded from the analysis. Crucially, this further ANOVA showed a significant three-way interaction, $F(1,29) = 18.66$, $MSE = .038$, $p < .001$, $\eta_p^2 = .49$ (See Figure 13). To examine differences in performance as a function of Context transition, a 2 Context of congruence (high congruent, low congruent) by 2 Congruence (congruent, incongruent) ANOVA was conducted for each type of transition, separately. Results revealed a significant two-way interaction for both conditions, whenever stimulus presentation followed a context repetition $F(1,19) = 11.67$, $MSE = .026$, $p < .003$, $\eta_p^2 = .38$ and when stimulus display switched between contexts $F(1,19) = 7.06$, $MSE = .013$, $p < .02$, $\eta_p^2 = .27$. The context specific proportion-congruent effect on the former condition was due to a greater magnitude of the congruence effect on the high proportion-congruent context (23.76%) than on the low proportion-congruent context (16.57%). The inverse pattern of results was observed in the condition where context of congruence switched between trials: a larger congruence effect on the low proportion-congruent context (23.33%) than on the high proportion-congruent context (17.86%). In terms of the magnitude of the congruence effects, these results replicate previous accounts (e.g., King et al., 2012) showing the relevance not only of the current context of congruence but of the context of congruence of the previous trial.

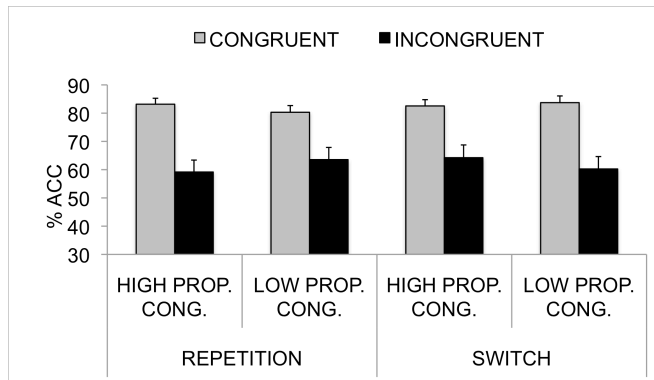


Figure 13. Percentage of correct responses as a function of Context transition, Context and Congruence. Error bars represent the standard error of the mean.

4.2.2. Questionnaire measures

Participants' reports revealed that 5 (25%) of the participants noticed a difference between the two contexts (e.g., "I made more errors on the upper half"). However, none of them reported being aware of the specific proportion-congruent manipulation, in accordance to previous studies (e.g., Crump et al., 2006; Heinemann et al., 2009; Sarmiento et al., 2012; King et al., 2012).

4.2.3. Surface ERP data

A time point-wise 2 Context of congruence (high congruent, low congruent) x 2 Congruence (congruent, incongruent) ANOVA on the Global Field Power data was conducted to identify modulations in the strength of the electric field at the scalp. There was a significant interaction between the two factors during two periods: ~163-227 ms and ~252-272 ms (see Figure 14). Crucially, including the factor Context transition revealed that shifting between contexts modulated the interaction between Congruence and Context of congruence only in the latter period, since a significant interaction between Context transition, Context of congruence, and Congruence was shown over the ~253-386 ms post-stimulus interval. We focus first on this earlier time period and then describe the results of the later one.

Post-hoc contrasts on the 163-227 ms period revealed a stronger response on congruent trials than incongruent trials in the high proportion-congruent context, whereas there was no statistical difference between congruent and incongruent trials in the low proportion-congruent context.

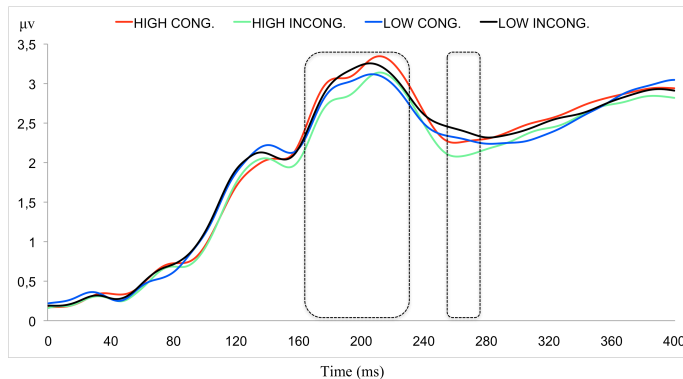


Figure 14. Group-average GFP waveforms and results of the ANOVA on the GFP as a function of time ($p < 0.05$; $> 12\text{ms}$) Dotted boxes represent intervals of statistical significant differences.

A hierarchical topographic cluster analysis was performed on the collective group-average ERPs to identify periods of stable electric field topography both within and between experimental conditions. 10 template maps that explained 93.59% of the variance were identified. At the group-average level two main results were observed (see Figure 15a). First, different topographic configurations over the ~163-227 ms post-stimulus period accounted for each condition as a function of its Congruence and the Context of congruence at which it was presented. Second, there was a faster onset for one of the templates maps as a function of the Context of congruence. This pattern was statistically verified in the single-subject ERPs using a spatial-correlation fitting procedure over the periods of interest (see Figure 15b). A 2 Context of congruence \times 2 Congruence \times 3 Map ANOVA revealed a significant interaction between Context and Map for the 163-227 ms period, $F(2,38) = 4.87$, $MSE = 367.2$, $p < 0.02$, $\eta_p^2 = .20$. The time that each map was presented was compared between contexts in order to analyse this interaction further.

RESULTS: DISCRIMINATION TASK

These post-hoc analyses showed that the green map predominated in responses in the high proportion-congruent context $F(1,19) = 8.42$, $MSE = 451.25$, $p < 0.01$, while there was a marginally significant effect for the blue map in the low proportion-congruent context $F(1,19) = 3.92$, $MSE = 255.61$, $p = 0.06$.

A three-way interaction between Context of congruence, Congruence and Map was also observed over the 163-127 ms period, $F(2,38) = 4.47$, $MSE = 52.5$, $p < 0.02$, $\eta_p^2 = .19$. As can be seen in Figure 15b, the green and blue template maps showed inverted patterns for the two contexts: the green map was present for a larger period in congruent than incongruent trials in the high proportion-congruent context, while the reverse was true for the low proportion-congruent context. The blue map showed the opposite pattern of results.

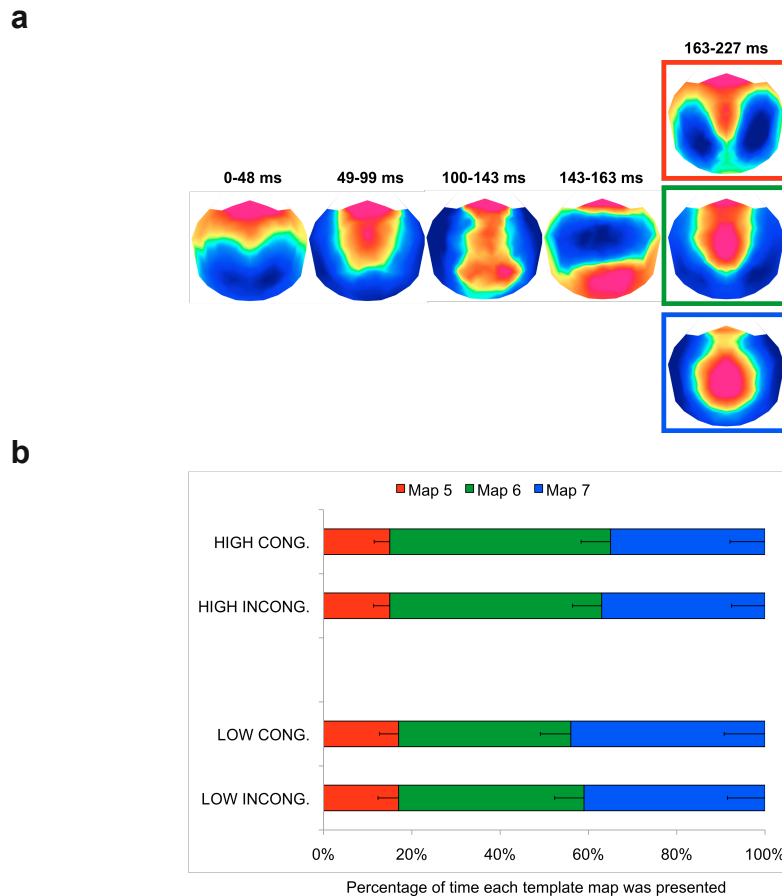


Figure 15. a Template maps from topographic cluster analysis. **b.** Results of the spatial correlation fitting in the single-subject data from each condition. Error bars represent the standard error of the mean.

As noted above, a point-wise Context transition (repetition, switch) x Context of congruence (high congruent, low congruent) x Congruence (congruent, incongruent) ANOVA on the global field power data showed a three-way interaction over a ~253-386 ms post-stimulus interval. To further understand this interaction each type of transition was, separately, subjected to a 2 Context of congruence (high congruent, low congruent) x 2 Congruence (congruent, incongruent) ANOVA, revealing that the interaction between Context of congruence and Congruence was statistically significant only after the repetition of the same context of congruence over the ~200 to 377 ms. Post-hoc comparisons

RESULTS: DISCRIMINATION TASK

showed that this interaction was driven by larger strength for the incongruent stimuli than the congruent stimuli on the low proportion-congruent context, while there were no difference between congruent and incongruent stimuli on the high proportion-congruent context (see Figure 16).

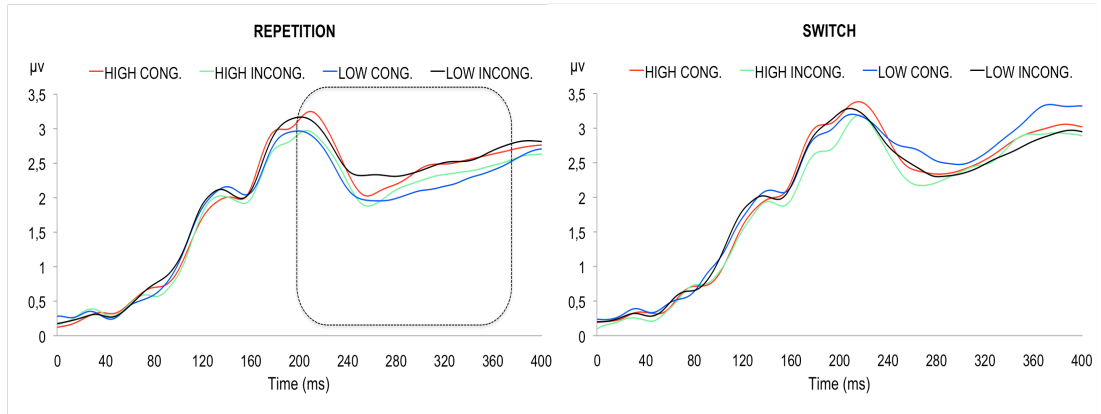


Figure 16. Group-average GFP waveforms and results of the ANOVA on the GFP as a function of time ($p < 0.05$; $> 12\text{ms}$) Dotted boxes represent intervals of statistical significant differences.

The results of the topographic cluster analysis identified 22 stable periods over the post-stimulus epoch that explained 95.72% of variance in the collective dataset. These template maps were then entered a fitting procedure on the single-subject data. The obtained values over the 253-386 ms period were submitted to a repeated measures ANOVA with 2 Context transition, 2 Context of congruence, 2 Congruence and 5 Map as within-subject factors. Results showed that the same map configuration accounted for the context specific proportion-congruent effect in both types of context transition (no significant four way interaction, $F < 1$).

4.2.4. Source estimations

Intracranial sources for the 163-227 ms period were estimated, as has been previously explained, to identify regions that were differently activated between congruent and incongruent condition as a function of the Context of congruence. Results from the 2 Context of congruence by 2 Congruence ANOVA revealed a significant interaction within

the temporal and frontal lobes (See Figure 17). Statistical contrast between congruent and incongruent source estimations within the high proportion-congruent context identified left superior (Max: -53, -28, 3 mm; BA 22) and left middle temporal (Max: -53, -29, 1 mm; BA 21) gyrus, as well as, left middle frontal gyrus (Max: -53, 2, 40 mm; BA 6) and left precentral gyrus (Max: -53, 1, 34 mm; BA 6) as the more activated regions during the processing of incongruent trials as compared to the congruent trials.

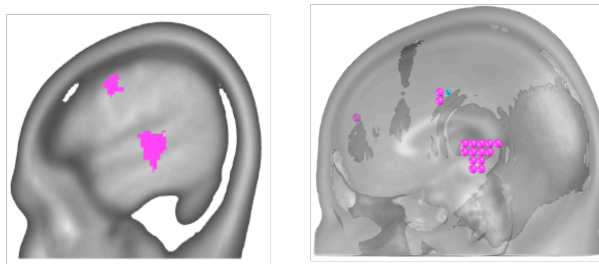


Figure 17. Results of the post-hoc comparisons between congruent and incongruent stimulation in the high proportion-congruent context over the 163-227 ms interval. Note that the positive values (warmer colors) indicate that the clusters are more active in the congruent condition than in the incongruent condition, while negative values (cooler colors) indicates the opposite.

Two main results arise from the analysis of the data from the multisensory discrimination task. First, the context specific proportion-congruent effect was observed over a first period (~163-227 ms) at all levels of analysis. This effect was shown at the global index of strength as a significant congruence effect only in the high proportion-congruent context, a context that was associated with the presence of a particular template map and with differently activated areas between congruent and incongruent stimuli. Second, the abovementioned effect was not affected by the context transition, since this variable modulated the context specific proportion-congruent effect over a later period (~253-386 ms), which probably accounted for the behavioural results more related to response-related processes.

5. Discussion

Visual duration discrimination can be affected by concurrent auditory stimulation (e.g., Klink, Montijn & van Wezel, 2011; Watkins, Shams, Josephs & Rees, 2007; Sekuler, Sekuler & Lau, 1997; for a general review see Shams & Kim, 2010). Sarmiento et al. (2012) recently showed that such audiovisual interactions depended on the particular context of congruence, with a larger congruence effect in a high proportion-congruent context than in a low proportion-congruent context. The aim of the present study was twofold: 1) to investigate whether the audiovisual stimuli used in Sarmiento et al.'s (2012) study resulted in larger brain response than the summed activity of the visual and auditory stimuli when presented in isolation; 2) to investigate the stage of stimulus processing at which specific context of congruence affected audiovisual interaction in the duration discrimination task. In the oddball task, auditory, visual and audiovisual (congruent and incongruent) stimuli were presented at two different locations. In the visual discrimination task audiovisual congruent and incongruent stimuli were displayed in these same locations that were associated with different proportions of congruence, giving rise to a high proportion-congruent context and a low proportion-congruent context.

Results of the oddball task revealed that responses to the AV condition differed in strength, topographical configuration and source activity when compared to the A+V condition. Specifically, superadditive effects occurred over the ~114-154 ms post-stimulus onset interval and were attributable to the modulation of scalp topographies as a result of different active generators for each condition. Greater activity for the AV than for the A+V was shown in the left cuneus and left middle occipital gyrus: areas related to audiovisual integration (e.g., Calvert et al., 2000, Dhamala et al., 2007; Cappe et al., 2010, 2012). These results suggest that the auditory and visual stimuli presented in this study and in our previous work (Sarmiento et al., 2012) were indeed integrated. There was no evidence that congruence modulated these interactions.

The analyses of the EEG data from the duration discrimination task revealed novel and interesting findings. First and contrary to the oddball task, the GFP analyses showed a reliable congruence effect (i.e., great strength for congruent than incongruent inputs) over the ~163-227 ms post-stimulus onset. Crucially, this congruence effect was only shown in the high proportion-congruent context. Changes at the topographic level supported the GFP analysis, since different topographies accounted for each of the context of congruence at the same stage of processing (~163-227 ms). It implied that neural areas were recruited in a different fashion at each context, even though the stimuli presented on both contexts were physically identical. Source estimations also reinforced the different stimulus processing between contexts: a group of fronto-temporal areas showed different activation between congruent and incongruent stimuli in the high proportion-congruent context, while no such differences in activation were observed in the low proportion-congruent context. Congruent and incongruent trials in the high proportion-congruent context differed in their scalp current densities at the primary auditory cortex (i.e., left superior and middle temporal cortex) the left middle frontal gyrus, and the primary motor cortex (i.e., left precentral gyrus). This activity pattern was lateralized to the left hemisphere, a result that could be explained by the particular nature of the task. Indeed, the left middle temporal and frontal gyri have been shown to be selectively activated for tasks related to the temporal properties of stimuli (e.g., Talati & Hirsch, 2005), and the left-hemisphere auditory cortex seems to respond specifically to rapid temporal stimulation (e.g., Belin et al., 1998; Zatorre & Belin, 2001).

The congruence effect was only observed on the discrimination task, where congruence was task-relevant in contrast to the oddball situation. Interestingly enough, the proportion of congruence modulated the congruence effect differently. We argue that the functional role of this modulation was to selectively strengthen the signal of the most reliable (and probable) information in order to obtain the most coherent and stable perceptual representation. Particularly, it reflected the influence of the auditory stimuli in

DISCUSSION

the processing of the visual stimulus only when the auditory information was reliable and matched that of the visual target in most of the occasions, that is, in the high proportion-congruent context. In the same way, the signal from the incongruent trials in this context (highly improbable and not reliable) was not strengthened (with respect to the other stimuli). On the other hand, congruent and incongruent audiovisual stimuli from the low proportion-congruent context were treated similarly given that auditory information was not reliable in the vast majority of the trials. It is worth noting here that the same number of congruent and incongruent stimuli was displayed throughout the experiment (and in every block of trials), and the only feature that categorized multisensory stimuli as more or less reliable was their location that was associated with a particular context of congruence.

King et al. (2012) in a previous fMRI study demonstrated that the context specific proportion-congruent effect was mediated by the type of transition between sequential contexts - repetition or switch of context from the current trial to the next trial -: the modulation of the congruence effect was present after the repetition of the same context. In the present study, the possibility of the influence of context transition over the abovementioned EEG proportion-congruent effect was ruled out, since the context repetition modulated the proportion-congruent effect at a later stage of processing (~253-386 ms). Therefore, the high temporal resolution of the EEG technique showed two different moments throughout the processing stream, which seem to be related to different processes. At a first window (~163-227 ms) the proportion-congruent effect was the outcome resulting from an on-line influence of the deployment of control over the perception of the relevant inputs in the current trial. Whereas, at a second period (~253-386 ms), the proportion-congruent effect was mediated by the type of transition and consequently the deployment of control over the performance was manifested in the subsequent trial, in line with King et al's (2012) findings. The electrophysiological effect over the second period might be related to decision-making and response selection processes and therefore probably responsible of the behavioral results.

Specifically, the behavioural results of the visual discrimination task revealed that the context specific proportion-congruent effect emerged when stimuli were repeatedly presented in the same context of congruence, and that this effect was inverted whenever a switch from one context to another occurred. In line with King, Korb and Egner (2012), we argue that stimulus location primed rapid control adjustments in order to deploy attentional resources adequately. Since the location of the target stimulus, and therefore the suitable control setting, could not be predicted before the stimulus onset itself, behavioral effects of the contextual cueing were only observed on the subsequent trial, once the specific contextual information had been retrieved.

In our view, selective attention was the mechanism responsible of the link between a particular location, a particular audiovisual stimulus (congruent or incongruent) and, crucially, a particular context of congruence (e.g., Treisman, 1996, 2006). Indeed, research in implicit learning has repeatedly shown that selective attention is critical in order to establish a link between predictive dimensions of a particular stimulus (probability of congruence and spatial location in this case), given that such association only emerges when these dimensions are relevant to the observer (e.g., Jiménez & Méndez, 1999; Crump, Vaquero & Milliken, 2008). In the present study, the location was crucial since the participant needed to localize the target (displayed in one of two possible locations) in order to perform the task. Once the association between proportion of congruence and location was established, different processing modes were implemented to perform the task as a function of the particular proportion of congruent and incongruent trials. As a result, the very same stimulus (e.g., congruent) were processed differently as a function of the location where it was presented (that determined the context of congruence).

Two important aspects of the results discussed so far should be taken in account. On one hand, the stimuli were identical in their physical properties in both contexts (i.e., same number of audiovisual congruent and audiovisual incongruent trials). This excludes the possibility that differences in brain responses were due to differences in stimulation or

DISCUSSION

stimulus-response associations for each context. On the other hand, the lack of awareness regarding the congruence manipulation by the participants is in accordance with previous findings using similar paradigms (e.g., Crump et al, 2008; Heinemann et al., 2009; King et al, 2012; Sarmiento et al., 2012). This implies that any effect of attention was unlikely under the voluntary control of participants.

In sum, the present study revealed that the brain did not deal with audiovisual stimuli always in the same manner. Instead, this depended on the particular context of congruence wherein they were presented that, in turn, was defined by their spatial location. Behaviourally, this resulted in modulation of the congruence effect as a function of the context in the current trial and in the preceding trial. Neurophysiologically, it resulted in a distinct pattern of activation as a function of the context of congruence, which reflects a different processing mode on each context responsible of the strengthening of the processing of the most reliable stimuli in each context, influencing decision-making and perhaps even the motor plan. We argue that selective attention would be responsible for this outcome, helping to associate each location with its proportion of congruence in first place and modulating the processing of the auditory and visual stimuli differently in each context in second place. This study sheds new light on the on-going debate regarding the interplay between attention and multisensory processing. Particularly, it suggests that indeed top-down control processes modulate the outcome of audiovisual inputs at different stages of the stimulus processing (i.e., see Mishra et al., 2010 for a related result) as a function of stimulus congruence (cf. Talsma et al., 2010). Importantly, the current results points to the fact that incongruity does not necessarily trigger top-down control, but that the relevance of the incongruity for the task, as well as the characteristics of the context (i.e., proportion of incongruent trials), determine the deployment of the particular set of control.

This study was supported by a Swiss National Science Foundation grant (310030B-133136) to Micah M. Murray, a student scholarship (AP2008-03662) from the Ministerio de Educación y Ciencia to Beatriz R. Sarmiento, grant SEJ2007-63645 from the Ministerio de Educación y Ciencia to Daniel Sanabria. We thank Diana Cutanda for her help with data collection.

CAPÍTULO 5. DISCUSIÓN

El trabajo que se ha realizado en esta tesis se ha centrado en explorar el papel que juega la atención en el procesamiento multisensorial. Concretamente, se ha pretendido aportar nueva evidencia en referencia a las condiciones que modulan la relación entre estos dos procesos, mediante la manipulación de la proporción de congruencia entre estímulos audiovisuales, característica que no ha sido estudiada previamente en el campo del procesamiento multisensorial.

Para la consecución de este objetivo se ha utilizado el paradigma de proporción de congruencia específico de contexto (CSPC) con una tarea de discriminación temporal audiovisual. Mediante la utilización de medidas conductuales y electrofisiológicas se ha observado que 1) tanto la tarea como los estímulos utilizados son susceptibles de dar lugar a efectos de integración multisensorial; 2) dicho efecto multisensorial se ve modulado por el contexto de proporción de congruencia en diferentes modos a lo largo del procesamiento de la información, desde los primeros momentos de procesamiento sensorial hasta los últimos momentos de selección y ejecución de la respuesta; 3) la congruencia entre los estímulos sólo manifiesta su influencia cuando es relevante para la tarea a realizar.

Con el objetivo de exponer lo más claramente posible los resultados obtenidos, así como sus implicaciones en el campo que nos ocupa, hemos organizado la discusión de esta tesis en los siguientes apartados. En una primera sección se destacan las aportaciones metodológicas que hacen de este trabajo un modo novedoso de estudiar tanto el procesamiento multisensorial, como el papel que la atención juega a la hora de modular dicho procesamiento. A continuación se hace una recapitulación de los principales resultados observados en los estudios llevados a cabo durante la realización de esta tesis. Para concluir, nos ocuparemos en otros dos apartados de argumentar las implicaciones que los estudios presentados tienen en el campo del procesamiento multisensorial, y en el marco de interacción entre atención y procesamiento multisensorial.

1. Aportaciones metodológicas

A nivel general tres son los aspectos novedosos a destacar en este trabajo. En primer lugar, la aproximación utilizada para medir el procesamiento multisensorial, que ha consistido en una tarea de discriminación temporal en la que se manipulaba la congruencia entre la duración de los estímulos visuales objetivos y los estímulos auditivos distractores. Con esta sencilla manipulación se pudo medir la diferencia en ejecución entre los ensayos en que los estímulos objetivos y distractores eran congruentes y los ensayos en los que eran incongruentes, obteniendo así la magnitud del efecto de congruencia para cada condición, o lo que es lo mismo, un índice directo de procesamiento multisensorial. Hasta donde nuestro conocimiento alcanza esta es la primera vez que se utiliza este índice en una tarea temporal de discriminación audiovisual. Si bien Klink y colaboradores (2011) encontraron un efecto similar al nuestro, el paradigma utilizado era completamente diferente ya que: 1) los participantes debían responder alternativamente a ambas modalidades sensoriales; 2) la duración de los estímulos variaba en un amplio rango; 3) la tarea consistía en comparar la duración entre dos estímulos; 4) median la ejecución en la tarea comparándola con la ejecución esperada a partir del índice psicofísico JND (del inglés *just noticeable difference*; diferencia mínima percibida entre dos estímulos) calculado individualmente para cada participante en una tarea previa.

El segundo punto a tener en cuenta es la utilización del paradigma de proporción de congruencia específica del contexto propuesto en 2006 por Crump y colaboradores. La relevancia del uso de este paradigma es triple. Por un lado, es la primera vez que se utiliza con estímulos multisensoriales, ya que en todos los estudios previos se manipulaba la congruencia entre dos dimensiones de la misma modalidad sensorial (v.gr., Crump y Milliken, 2009; Heinemann, Kunde, y Kiesel, 2009; King, Korb, y Egner, 2012). Por otro lado, hasta el momento las investigaciones que trataban de estudiar el papel de la atención en el procesamiento multisensorial no se habían planteado manipular la

proporción de congruencia, sino que basaban sus manipulaciones en comparar la ejecución en condiciones de atención dividida entre modalidades frente a condiciones de atención focalizada en una sola modalidad (v.gr., Larse, McIlhagga, Baert y Bundesen, 2003; de Jong, Toffanin, y Harbers, 2010); o en comprobar si una señal en una modalidad sensorial podía influir en el procesamiento de otra modalidad sensorial (véase Spence y Driver, 2004); o en observar si prestar atención a una localización espacial modula el procesamiento de estímulos multisensorial presentados en dicha localización frente a estímulos presentados en otra (v.gr., Talsma y Woldorff, 2005), etc. Por último, la creación de dos contextos en los que la proporción de congruencia fue manipulada, permitió estudiar no sólo como los estímulos multisensoriales se comportaban ante una situación en la que se da incongruencia entre modalidades, sino que da un paso más al permitir observar como se procesaban estos estímulos audiovisuales cuando se presentaban en una situación de incongruencia casi constante, frente a cuando se presentaban en una situación generalmente congruente.

El tercer aspecto que merece ser mencionado es la técnica de registro utilizada, así como el método de análisis. Hasta el momento ningún estudio había combinado el paradigma CSPC con el registro de actividad electroencefalográfica. Dicha combinación permitió conocer el curso temporal junto con los correlatos neurales que subyacen a este efecto. Si bien algunas investigaciones centradas en el estudio del procesamiento multisensorial han utilizado la técnica de electrical neuroimaging (v.gr., Thelen, Cappe y Murray, 2012), ésta no ha sido tomada en cuenta en las investigaciones que han estudiado el papel de la atención en el procesamiento multisensorial. La importancia del análisis que se ha llevado a cabo en los trabajos de esta tesis es la aportación de información complementaria a la que ya se había obtenido a través de otras técnicas de neuroimagen.

2. Resumen de los resultados principales

El primer resultado, que constituye el germen del que parten todos los estudios de esta tesis, se resume en la siguiente frase: *la duración percibida de un estímulo visual depende de la duración real de un estímulo auditivo sincrónicamente presentado.*

Este efecto obtenido en el primero de los estudios comportamentales indicaba que la información auditiva, aunque irrelevante, estaba influyendo en el procesamiento de la información visual, tanto a nivel perceptivo (d') como a nivel de criterio de respuesta (c). Obteníamos así un efecto de interacción audiovisual sobre el que poder estudiar la influencia de la manipulación de la proporción de congruencia.

En el segundo de los estudios se aplicó dicha manipulación en bloques separados de alta y baja proporción de congruencia. Los resultados mostraron que el efecto de interacción audiovisual era menor en el contexto de baja proporción de congruencia, apuntando así a que el procesamiento de estímulos audiovisuales podía ser modulado por la frecuencia de ensayos congruentes e incongruentes, cuando la presentación de dichos contextos se hacía en bloques separados.

El siguiente paso fue comprobar si esta misma modulación del procesamiento multisensorial se podía obtener cuando la presentación de los contextos de congruencia se intercalaba dentro del mismo bloque de ensayos. Para ello, en el Experimento 3 se asoció cada uno de los dos contextos de congruencia con una de dos localizaciones en la pantalla del ordenador (i.e., encima o debajo del punto de fijación). Los resultados de este experimento no mostraron una modulación del efecto de interacción audiovisual en función del contexto de congruencia.

En el Experimento 4 se remarcó la existencia de dos contextos diferenciados mediante la presentación de una larga línea horizontal en el punto de fijación, al tiempo que las localizaciones en que se presentaban los estímulos de uno y otro contexto eran más distantes entre sí. Añadiendo estas modificaciones, se observó que el efecto de interacción audiovisual era menor en el contexto de baja proporción de congruencia,

130

acorde a lo encontrado en el Experimento 2, pero en este último estudio cuando ambos contextos se presentaban intercalados. Los análisis basados en la teoría de detección de señales mostraron que el efecto de proporción de congruencia asociado al contexto se relacionaba con un incremento en la sensibilidad perceptiva (d') para el contexto de baja proporción de congruencia.

Las medidas subjetivas recogidas mediante cuestionario en los Experimentos 2, 3 y 4, mostraron que los participantes no eran conscientes de la manipulación de congruencia que se había llevado a cabo, aún cuando estaban siendo sensibles a ella y utilizándola para la realización de la tarea.

La Serie Experimental I puso de manifiesto que el procesamiento de estímulos audiovisuales se veía modulado por procesos atencionales derivados de la manipulación de la proporción de congruencia. Esta modulación ocurría tanto cuando los contextos de proporción de congruencia se presentaban de manera consecutiva en bloques separados, como cuando se presentaban intercalados dentro del mismo bloque. El despliegue de dichos procesos atencionales de selección de información era independiente del conocimiento explícito de los participantes sobre la manipulación de la proporción de congruencia. Los efectos encontrados parecen relacionarse con una mejora en la percepción de los estímulos en el contexto de baja proporción de congruencia, al menos cuando la presentación de los estímulos en ambos contextos se entremezclaba en el mismo bloque de ensayos.

La Serie Experimental II centrada en el registro de actividad electroencefalográfica pretendía solventar las limitaciones del estudio comportamental. En primer lugar, comprobar si el tipo de procesamiento multisensorial que subyacía al efecto de congruencia audiovisual encontrado - i.e., la duración percibida de un estímulo visual depende de la duración real de un estímulo auditivo sincrónicamente presentado - podía ser catalogado como integración multisensorial en sentido estricto, o si por el contrario se trataba de una interacción entre modalidades sensoriales que influía en el percepto final,

RESUMEN DE LOS RESULTADOS PRINCIPALES

pero que no creaba una experiencia perceptiva cualitativamente diferente de la que se derivaría de la percepción aislada de estímulos visuales y auditivos. En segundo lugar, aclarar en qué momento del procesamiento estaba teniendo lugar la influencia de los procesos asociados a la manipulación del contexto de congruencia, gracias a la gran resolución temporal que caracteriza a la técnica de registro electrofisiológico, complementando las conclusiones obtenidas en la Serie Experimental I basadas en la aplicación de la teoría de detección de señales.

En esta serie experimental, junto con la tarea de discriminación temporal, se añadió una tarea oddball con el objetivo de identificar el tipo de procesamiento multisensorial que estaba teniendo lugar. A continuación se presentan los resultados obtenidos para cada una de las tareas.

Los resultados de la tarea oddball mostraron un efecto de integración multisensorial alrededor de los 115-154 ms tras la presentación de los estímulos. Concretamente, se observaron diferencias entre los estímulos audiovisuales y la suma de los estímulos auditivos y visuales en potencia, en topografía y actividad cerebral de fuentes neurales occipitales que se asocian con procesos de integración multisensorial (i.e., cúneo izquierdo y giro occipital medial). Por lo tanto, estos datos ponen de manifiesto que el tipo de procesamiento que se está llevando a cabo da lugar a la integración de los estímulos audiovisuales presentados.

La tarea de discriminación temporal mostró a nivel conductual como los estímulos audiovisuales se procesaron de manera diferente en función de la proporción de congruencia asociada a la localización en que se presentaban dichos estímulos. Además de la proporción de congruencia, también se puso de manifiesto la importancia del tipo de contexto previo, de manera que el efecto de congruencia era menor en el contexto de baja proporción de congruencia sólo cuando el ensayo previo y el ensayo actual pertenecieron al mismo tipo de contexto. Sin embargo, cuando se dio un cambio de contexto entre

ensayo previo y actual la magnitud del efecto de congruencia era menor para el contexto de alta proporción de congruencia.

La actividad electroencefalográfica registrada durante la tarea de discriminación mostró la influencia del contexto de proporción de congruencia sobre los efectos de procesamiento multisensorial en distintos momentos del procesamiento de la información.

El primer momento en el que se observó una modulación de la actividad eléctrica asociada al tipo de ensayo (i.e., congruente o incongruente) en función del contexto de congruencia fue alrededor de los 163 ms durando aproximadamente hasta los 227 ms. Durante este periodo una mayor potencia se registró para los ensayos congruentes en relación a los incongruentes en el contexto de alta proporción de congruencia. Esta diferencia estuvo acompañada de la presencia de distintas topografías para los ensayos congruentes e incongruentes, derivada de la mayor actividad de áreas frontotemporales (i.e., giro temporal medial y superior; giro frontal medial y giro precentral del hemisferio izquierdo) en relación al procesamiento de los ensayos incongruentes. El efecto de proporción de congruencia específico del contexto observado en este primer momento era un efecto puro, es decir, no estaba influenciado por el tipo de contexto que se había presentado previamente. Este tipo de influencia de la transición entre contextos se puso de manifiesto en momentos más tardíos del procesamiento de la información (i.e., ~253-386 ms), al mostrar que la proporción de congruencia sólo modulaba el resultado del procesamiento multisensorial cuando se daba una repetición de contexto del ensayo previo al ensayo actual, mostrando un incremento en la señal asociada a los ensayos incongruentes del contexto de baja proporción de congruencia, en línea con los efectos mostrados a nivel comportamental.

Replicando los resultados de la Serie Experimental I, la puesta en marcha de los mecanismos cognitivos responsables de selección de información relevante en cada contexto, se llevó a cabo independientemente del conocimiento consciente de la manipulación de la congruencia por parte de los participantes.

3. Integración audiovisual

A la hora de estudiar el procesamiento multisensorial en los trabajos de esta tesis hemos optado por un enfoque en el que se manipula la congruencia de una de las características físicas de los estímulos, concretamente la duración. Esta elección ha permitido aportar conocimiento sobre como la duración de estímulos en una modalidad sensorial influye sobre otra modalidad, campo que no ha sido ampliamente estudiado, mientras que sobre otras características de los estímulos como la localización o la asincronía posemos un conocimiento más asentado (v.gr., Alais y Burr, 2004; Navarra, Vatakis, Zampini, Soto-Faraco, Humphreys y Spence, 2005, respectivamente).

Es importante destacar que aunque la característica central de nuestros estudios pertenece a la dimensión temporal, las investigaciones no se circunscriben al campo de la percepción del tiempo, referida ésta al estudio de la percepción subjetiva de la duración de los eventos, como la estimación del tiempo que transcurre entre la presentación de dos estímulos (v.gr., Eagleman, Tse, Buonomano, Janssen, Nobre y Holcombe, 2005; Correa, Lupiáñez y Tudela, 2006). Aquí tomamos la duración como una característica que define a un objeto, exactamente igual que lo definiría el color o el tamaño, y manipulamos su congruencia entre modalidades sensoriales con el objetivo de comprender como el sistema cognitivo resuelve la incongruencia, y no con el objetivo de conocer como un reloj interno o un marcapasos da lugar a la percepción que tenemos del transcurso del tiempo.

El estudio llevado a cabo muestra como la duración percibida de un estímulo visual depende de la duración real de un estímulo auditivo irrelevante cuando la tarea consiste en discriminar la duración del estímulo visual. Este efecto multisensorial va en concordancia con estudios previos que ya habían mostrado la influencia que la modalidad auditiva tiene sobre la modalidad visual cuando lo que se juzga es alguna característica que pertenece a la dimensión temporal (véase Shams y cols., 2000 para juicios de numerosidad; Morein-Zamir y cols., 2003 para juicios de orden temporal; Klink y cols., 2011 para juicios de duración). Este hecho se debería a la mayor resolución temporal que

tiene la modalidad auditiva sobre la modalidad visual, como ya hemos comentado a lo largo de esta tesis. Teniendo esto en cuenta, nos planteamos la cuestión de que estemos ante un efecto asimétrico, es decir, que si lo que hubiese que juzgar fuese la duración del estímulo auditivo, la presentación de un estímulo visual irrelevante no ejercería ninguna influencia sobre el procesamiento del estímulo auditivo. Apoyarían esta hipótesis estudios como el de Klink y colaboradores (2011) que evaluando cual de dos estímulos tiene mayor duración, observaron que cuando se preguntaba por la modalidad visual los juicios de duración se veían modificados por la duración de un estímulo auditivo irrelevante, mientras que esto no ocurría cuando el estímulo que se debía juzgar era el auditivo acompañado de un estímulo visual (véase también Chen y Yeh, 2009). También Andersen, Tiippana y Sams (2004) observaron que la presentación de varios sonidos (beeps en este caso) modificaba la percepción del número de estímulos visuales presentados. Sin embargo, esto no ocurría a la inversa, excepto en el caso en que la señal auditiva perdía calidad (i.e., frecuencia cercana al umbral mínimo de audición), de manera que la modalidad visual ganaba peso volviéndose más fiable, como se predecía a partir del modelo MLE (véase Ernst y Banks, 2002) presentado en la introducción.

El efecto de duración audiovisual aquí registrado se midió en un primer momento a nivel conductual calculando el índice de congruencia (i.e., diferencia entre los ensayos congruentes e incongruentes), lo que permitió identificar la modulación que una modalidad sensorial estaba ejerciendo sobre la otra. Sin embargo, la pregunta que nosotros queríamos responder era si dicho efecto constituía la integración de la información de ambas modalidades sensoriales. Los análisis basados en la teoría de detección de señales aportaron resultados ambiguos ya que mostraban una mayor sensibilidad perceptiva ante la presentación de estímulos unimodales frente a la presentación de estímulos audiovisuales, y no mostraban diferencias entre los estímulos audiovisuales congruentes e incongruentes. Lo que hubiésemos esperado sería una mayor sensibilidad perceptiva ante los estímulos audiovisuales congruentes frente a los unimodales, debido a

la ventaja ofrecida por la redundancia de la señal. Al igual que diferencias entre los ensayos congruentes e incongruentes, puesto que en el segundo caso la información auditiva interfiere con el procesamiento visual, de modo que se podría esperar una menor sensibilidad perceptiva, o incluso mayor si el sistema primase el procesamiento de la información visual para contrarrestar el efecto de la información auditiva incongruente. Además, también se encontró que el criterio de respuesta se veía modificado en función de la congruencia de los estímulos, lo que indicaría un sesgo a la hora de seleccionar y emitir la respuesta, pero no a la hora de integrar la información. La diferente influencia que la congruencia está teniendo a nivel perceptivo y decisional va en la línea de estudios como el de Sanabria y colaboradores (2007b) quienes obtuvieron el mismo patrón de resultados aquí presentados, utilizando un paradigma con estímulos dinámicos tanto en una tarea go/no-go como en una tarea de discriminación. Ambos estudios basan sus resultados e infieren sus conclusiones de las premisas impuestas desde teoría de detección de señales, lo que nos lleva a ser cautelosos a la hora de interpretar estos resultados dentro del campo del procesamiento multisensorial.

El siguiente paso fue buscar un método complementario para estudiar el nivel al cual se estaba dando la modulación del procesamiento audiovisual, lo que desembocó en el registro de la actividad electroencefalográfica, para así poder medir el efecto utilizando el modelo aditivo - i.e., $AV \neq A+V$ - desarrollado desde el área del procesamiento multisensorial.

Este estudio mostró un efecto de superaditividad, desde los 114 a los 154 ms, que estaría indicando la integración de los estímulos audiovisuales (véase Stein y Meredith, 1993; Giard y Peronet, 1999; sin embargo véase Holmes, 2009, para una interpretación alternativa en términos de artefacto estadístico). Es habitual que los efectos de integración comiencen un poco antes en el tiempo (i.e., alrededor de los 80 ms tras la presentación del estímulo). Teniendo en cuenta esta temporalidad se había modificado la duración de los estímulos con respecto a la Serie Experimental I, con el objetivo de que posibles

efectos debidos a la disparidad en la duración de los estímulos en la condición incongruente no se confundiesen con los efectos de la integración audiovisual. Sin embargo, el intervalo temporal en el que se observó el efecto puede llevar a una interpretación en términos de disparidad, es decir, el estímulo más corto terminaba a los 100 ms y a partir de ese momento se detectaba una diferencia entre estímulos congruentes e incongruentes, y eso es lo que la actividad eléctrica estaba registrando. Sin embargo, algunas evidencias nos llevan a descartar esta interpretación. Por un lado, el efecto de superadividad obtenido. Por otro lado, el hecho de que no hubiese un efecto de congruencia, es decir que el sistema no detectase una diferencia entre ensayos congruentes e incongruentes cuando la congruencia no era relevante (i.e., tarea oddball en la que no había que responder a la congruencia). En tercer lugar, otras investigaciones han mostrado efectos de integración multisensorial en intervalos temporales similares a los de nuestro estudio. Por ejemplo, Mishra y colaboradores (2007) observaron un primer efecto de integración multisensorial en el intervalo de ~100-130 ms tras la presentación de la estimulación, e incluso otros estudios como el llevado a cabo por Raij y colaboradores (2000) presentaban ventanas más tardías (v.gr., ~225 ms). Por último, las fuentes cerebrales que subyacen a los efectos encontrados son zonas de integración multisensorial (véase Calvert y cols., 2000; Dhamala y cols., 2007) situadas en áreas corticales de bajo nivel, en línea con estudios que muestran modulación de áreas visuales primarias ante la presentación de información auditiva concurrente (véase Watkins, Shams, Josephs y Rees, 2007).

Tomando en cuenta la totalidad de estas pruebas podemos defender que el efecto encontrado constituye la integración de estímulos auditivos y visuales. Si bien no descartamos la posibilidad de que haya efectos más tempranos a los que o bien el paradigma, o bien las técnicas de registro y análisis utilizadas no hayan sido sensibles.

En conclusión, presentamos en esta tesis una tarea de discriminación temporal audiovisual que constituye una nueva y sencilla aproximación al estudio de la integración

multisensorial. Concretamente, la medición a nivel conductual de la magnitud del efecto de congruencia nos permite hacer inferencias sobre el procesamiento multisensorial apoyándose en la evidencia electrofisiológica de integración.

4. Integración audiovisual y mecanismos de selección de información

El objetivo central de esta tesis era profundizar en el estudio de la interacción que se da entre el procesamiento multisensorial y los procesos atencionales. Como acabamos de comprobar el efecto de integración multisensorial estaba definido por la tarea de duración audiovisual. Así que aplicamos el paradigma de proporción de congruencia de contexto a este efecto audiovisual.

Recordemos que dicho paradigma implica procesos de atención selectiva en dos sentidos. Por una parte, es necesario que se preste atención a las características relevantes de la tarea para que se pueda aprender la relación entre ellas - en este caso la localización y la proporción de congruencia -, en línea con la función de la atención que se encarga de integrar las características relevantes para la conjunción de las unidades apropiadas. Por otro lado, en función de la proporción de congruencia se debe poner en funcionamiento un mecanismo de selección de información diferente - atender a estímulos visuales y auditivos o atender sólo a estímulos visuales -, lo que se relaciona con la segunda de las funciones de la atención destacadas en esta tesis, es decir el sesgo del procesamiento de la información hacia las características relevantes dependiendo del contexto concreto.

El primer resultado interesante es el hecho de que la codificación de la congruencia sólo tiene lugar cuando ésta es relevante para el objetivo a conseguir. Así, vemos como en la tarea de duración se observa un efecto de congruencia, no sólo a nivel conductual, sino también a nivel de actividad electrofisiológica, efecto que no se mostró en la tarea oddball en la que la duración de los estímulos era totalmente irrelevante. Este resultado

apoya la idea de que se está prestando atención a la congruencia entre los estímulos lo que permitirá asociarla como una característica más del evento. Esta conclusión está en línea con la idea de Treisman (1998, 2006) que plantea que la atención es el mecanismo encargado de agrupar en una unidad las características que pertenecen a un mismo evento. Aproximaciones similares a esta serían la propuesta desde el ámbito de la memoria por Crump y colaboradores (2009), quienes defienden que la atención selectiva es el mecanismo encargado de unificar en una representación episódica toda la información del evento; o desde la perspectiva del aprendizaje asociativo, donde Verguts y Notebaert (2009) proponen que las representaciones que están activas en un momento determinado, que son las relevantes para la tarea, se asocian a partir del fortalecimiento de las conexiones entre las áreas cerebrales implicadas.

En el caso que nos ocupa, el mecanismo de atención selectiva apropiado para cada uno de los contextos también es una característica del evento que, por lo tanto, conforma la unidad. De este modo, cuando el estímulo se presenta se activan todas las características asociadas a él, incluido el modo de procesar la información relevante e irrelevante. Dado que los eventos comparten todas sus características a excepción de la localización y el set de procesamiento adecuado, una vez que se conoce la localización del estímulo - i.e., tras su presentación -, rápidamente se activan el resto de características del evento, incluido el tipo de procesamiento que se debe llevar a cabo. Será este procesamiento de arriba-abajo el responsable de los efectos de congruencia diferentes entre contextos.

La cuestión que nos planteamos a continuación es demostrar como nuestros estudios apoyan la influencia de la atención sobre el procesamiento de los estímulos audiovisuales acorde al esquema que acabamos de comentar.

Los resultados conductuales de nuestros estudios muestran una adaptación del procesamiento audiovisual en función de la proporción de ensayos congruentes e incongruentes, es decir, un mayor efecto de congruencia para el contexto de alta

proporción de congruencia. En la Serie Experimental I, este efecto se muestra particularmente claro en el Experimento 2, debido a una mejor ejecución en los ensayos congruentes en el contexto caracterizado por una alta proporción de ensayos congruentes, y a una mejor ejecución en los ensayos incongruentes cuando el contexto se caracteriza por una mayor cantidad de ensayos incongruentes. Este efecto es menos potente en el Experimento 4 - donde solamente se da una pequeña mejora en la ejecución de los ensayos incongruentes en el contexto de baja proporción de congruencia -, llegando a no ser observable en los estudios de la Serie Experimental II.

La diferencia crítica entre estos estudios está en la forma de presentación de los contextos. En el Experimento 2 de la Serie Experimental I los contextos se presentan independientemente uno del otro en bloques separados, mientras que en el resto de experimentos los contextos se presentan entremezclados dentro del mismo bloque de ensayos. Lo que defendemos que está sucediendo, es que en el experimento bloqueado se pone en funcionamiento un modo de selección de la información que permanece constante a lo largo de todo el bloque, mientras en los otros estudios el set atencional debe variar en cada ensayo en función del tipo de contexto (i.e., localización). De manera que en esta última condición el tipo de transición entre contextos (i.e., repetición o cambio) influye en el resultado final, cosa que no ocurre en el Experimento 2 donde constantemente se da una repetición de contexto. Muestra de ello es que cuando tenemos en consideración la transición del contexto a la hora de llevar a cabo los análisis estadísticos obtenemos el efecto CSPC esperado (i.e., mayor efecto de congruencia en el contexto de alta proporción de congruencia) cuando se repite el tipo de contexto en ensayos secuenciales, que sería la condición que tenemos en el Experimento 2; y obtenemos el patrón invertido cuando se da un cambio de contexto (el Experimento 4 de la Serie Experimental I, muestra el efecto en términos de magnitud para este contexto, aunque no se observa un efecto estadístico significativo).

Lo que podemos inferir de estos resultados es que el procesamiento de los estímulos multisensoriales se puede modular por mecanismos atencionales de selección. Dichos mecanismos se despliegan de manera rápida y flexible en función de las necesidades particulares de cada contexto, ya que es necesario esperar a que aparezca el estímulo para poner en marcha dichos mecanismos, pues no es posible anticipar el contexto al que habrá que contestar. Destacando que la puesta en marcha de dichos mecanismos aunque es rápida, no es inmediata, necesitando al menos un ensayo para poder ejercer su influencia. Este dato va en línea con los estudios de Crump y colaboradores, aunque ellos defiende que la puesta en marcha de los mecanismos atencionales es mucho más inmediata, ocurriendo en el mismo ensayo. Esta diferencia puede ser debida al tipo de tarea utilizada ya que ellos manipulan la congruencia entre estímulos visuales que se presentan secuencialmente - i.e., la presentación central de una palabra que indica el nombre de un color va seguida de la presentación de la figura de color en una de las dos posibles localizaciones -, y observan como esto influye en los mecanismos de resolución de una tarea Stroop, mientras en nuestro estudio la presentación de los estímulos es simultánea - i.e., el estímulo visual y auditivo se presentan al mismo tiempo en uno de los dos contextos - (véase King y cols., 2012; Spape y Hommel, 2008 para aproximaciones similares). Sería interesante conocer los resultados de sus tareas teniendo en cuenta el tipo de transición de contexto para poder argumentar sobre los factores que determinan la puesta en marcha de los mecanismos de selección de la información.

Los resultados obtenidos a nivel comportamental también los encontramos a nivel electroencefalográfico. Alrededor de los 253-386 ms observamos el efecto de congruencia en el contexto de baja proporción de congruencia únicamente tras la repetición del contexto previo. En concreto, la respuesta neural de los estímulos incongruentes en este contexto presenta un incremento en su potencia. Este dato concuerda con la mejora en la ejecución en los ensayos incongruentes del contexto de baja proporción de congruencia que se observa en el Experimento 4 de la Serie Experimental I, así como en la Serie

Experimental II para la condición en que se repite el contexto entre el ensayo previo y el actual. Probablemente esta potenciación de la señal para los ensayos incongruentes, sea la responsable del aumento en la sensibilidad perceptiva (d') en el contexto de baja proporción de congruencia que se obtuvo en el Experimento 4. Además, estos datos añaden información a la idea propuesta por King y colaboradores (2012) al mostrar que la influencia de los procesos atencionales en función del tipo de transición entre contextos se da en estadios del procesamiento asociados a selección y ejecución de la respuesta, información que, debido a la baja resolución temporal de la técnica de IRMf, no pudo ser observada en su estudio, pero si en el nuestro mediante la técnica de registro electroencefalográfico.

Hasta ahora hemos obtenido evidencia convergente a nivel conductual y neural de la influencia de procesos atencionales sobre el procesamiento multisensorial. Este efecto está relacionado con procesos tardíos (253-386 ms), lo que llevaría a concluir que la atención se ocupa de sesgar la selección de la respuesta, pero no influye a la hora de integrar los estímulos en un único percepto.

Sin embargo, nuestros estudios muestran que también a nivel temprano la atención está teniendo un efecto sobre el procesamiento multisensorial. Se observó el efecto de proporción de congruencia específico del contexto entre los ~163 y los ~227 ms, guiado por una mayor potenciación de la señal en los estímulos congruentes del contexto de alta proporción de congruencia, mientras no se observó dicho efecto de congruencia en el contexto de baja proporción. Parece pues, que la localización del estímulo activa mecanismos diferentes para el contexto de alta proporción y el de baja proporción de congruencia, que relacionamos con la relevancia de la estimulación auditiva en cada uno de los contextos. Así, la potenciación de la señal de los estímulos congruentes pertenecientes al contexto de alta proporción de congruencia está poniendo de manifiesto la influencia de la estimulación auditiva, que en esta condición aporta información fiable, cosa que no ocurre en los estímulos incongruentes, ni en el contexto de baja proporción

de congruencia. Apoya esta hipótesis el hecho de que diferentes patrones de actividad neural caracterizan a cada uno de los contextos. Observándose diferencias en la actividad neural de las áreas auditivas primarias entre los estímulos del contexto de alta proporción de congruencia, al tiempo que no se presentaron diferencias en el contexto de baja proporción de congruencia. Este dato indicaría que en la baja proporción de congruencia, debido a la alta probabilidad de conflicto, se opta por prestar atención selectivamente a la modalidad visual, inhibiendo así la actividad de las áreas auditivas (veáse Jonhson y Zatorre, 2005).

Podemos concluir por tanto, que el efecto de integración multisensorial observado es susceptible de modularse por procesos atencionales derivados de la manipulación de los contextos de proporción de congruencia en distintos momentos del procesamiento. Estos procesos atencionales se activan de un modo rápido, flexible e independiente del conocimiento explícito que se tenga sobre la manipulación de la congruencia.

Partiendo del modelo de interacción polifacética de la atención y el procesamiento multisensorial propuesto por Talsma y colaboradores (2010) observamos como procesos atencionales de arriba-abajo son requeridos en condiciones de conflicto entre las representaciones de los estímulos. Además añadimos dos elementos importantes que completan y delimitan el modelo. Primero, el hecho de que no cualquier clase de conflicto desata la puesta en marcha de los mecanismo de atención selectiva, sino que la competición se debe dar entre dimensiones que son relevantes para la meta concreta. Así nuestros datos muestran como la congruencia sólo produce un efecto cuando es relevante para la tarea, es decir, cuando es atendida e integrada como parte del evento, lo que ocurre en la tarea de discriminación de la duración, pero no en la tarea oddball. Sería interesante realizar un estudio en el que la tarea oddball se presente acompañada de la manipulación de los contextos de congruencia, para corroborar que no sólo no tendríamos el efecto de congruencia, sino que tampoco se daría el efecto de proporción de congruencia. Segundo, no es simplemente el conflicto *per se* el que pone en marcha los

procesos de selección de respuesta, sino que la relevancia de dicha incongruencia es la que guía el despliegue de dichos mecanismos. Dicho en otro modo, no toda incongruencia o competición entre estímulos es resuelta del mismo modo o utilizando los mismos mecanismos atencionales, sino que el tipo de sesgo atencional dependerá de cuán común sea la incongruencia en la situación concreta. Esto se pone de manifiesto en nuestros estudios en los resultados que muestran como los mismos estímulos incongruentes disparan mecanismos atencionales diferentes en función de que se presenten en un contexto donde la incongruencia es la norma, o en un contexto donde rara vez se presenta incongruencia.

Los trabajos realizados a lo largo de esta tesis doctoral apuntan a que el sistema cognitivo es sensible al conflicto entre las distintas modalidades sensoriales, siempre y cuando dicho conflicto sea relevante para la tarea a ejecutar. Además nuestro sistema cognitivo es capaz de computar la probabilidad de conflicto y asociarla con los mecanismos de selección necesarios para resolver dicha incongruencia, generando de este modo una percepción coherente del mundo. Tanto la asociación de las características en un único evento, como la puesta en marcha de los mecanismos se lleva a cabo sin la necesidad de un conocimiento consciente por parte de la persona, y de un modo altamente flexible.

REFERENCIAS BIBLIOGRÁFICAS

- Alais, D. y Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration, *Current Biology*, 14, 257-262.
- Alink, A., Singer, W. y Muckli, L. (2008). Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *The Journal of Neuroscience*, 28(11), 2690-2697.
- Altieri, N., Stevenson, R.A., Wallace, M.T. y Wenger, M. J. (2013). Learning to associate auditory and visual stimuli: Behavioral and neural mechanisms. *Brain Topography*. Published online 26 November 2013.
- Alvarado, J.C., Vaughan, J.W., Stanford, T.R., y Stein, B.E. (2007). Multisensory versus unisensory integration: contrasting modes in the superior colliculus. *Journal of Neurophysiology*, 97, 3193-3205.
- Andersen, T.S., Tiippana, K. y Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, 21, 301-308.
- Appelbaum, L.G., Smith, D.V., Boehler, C.N., Chen, W.D., y Woldorff, M.G. (2011). Rapid modulation of sensory processing induced by stimulus conflict. *Journal of Cognitive Neuroscience*, 23(9), 2620-2628.
- Aschersleben, G., Bachmann, T., y Müsseler, J. (Eds.) (1999). *Cognitive contributions to the perception of spatial and temporal events*. Amsterdam: Elsevier/North-Holland.
- Avillac, M., Hamed, S.B. y Duhamel, J.R. (2007). Multisensory Integration in the Ventral Intraparietal Area of the Macaque Monkey. *The Journal of Neuroscience*, 27(8), 1922-1932.
- Banati, R.B., Goerres, G.W., Tjoa, C., Aggleton, J.P. y Grasby, P. (2000). The functional anatomy of visual-tactile integration in man: a study using positron emission tomography. *Neuropsychologia*, 38, 115-124.
- Barracough, N.E., Xiao, D., Baker, C.I., Oram, M.W. y Perrett, D. I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *Journal of Cognitive Neuroscience*, 17, 377-391.
- Battaglia, P.W., Jacobs, R.A. y Aslin, R.N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America*, 20(7), 1391-1397.
- Beauchamp, M.S., Lee, K.E., Argall, B.D., y Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron*, 41, 809-823.
- Beck, D.M. y Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature Neuroscience*, 8 (8), 1110-1116.

REFERENCIAS BIBLIOGRÁFICAS

- Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M.C. y Samson, Y. (1998). Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience*, 10, 536-540.
- Bertelson, P., Vroomen, J., De Gelder, B. y Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, 62 (2), 321-332.
- Vroomen, J., Bertelson, P. y De Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & Psychophysics*, 63(4), 651-659.
- Bertelson P. y Aschersleben G. (2003). Temporal ventriloquism: Cross-modal interaction on the time dimension: 1 Evidence from auditory-visual temporal order judgment. *International Journal of Psychophysiology*, 50, 147-155.
- Bertelson, P., y de Gelder, B. (2004). The psychology of multimodal perception. En C. Spence y J. Driver (Eds.). *Crossmodal space and crossmodal attention*. Oxford, UK: Oxford University Press.
- Bieser, A. (1998). Processing of twitter-call fundamental frequencies in insula and auditory cortex of squirrel monkeys. *Experimental Brain Research*, 122 (2), 139-148.
- Botvinick, M., Cohen, J.D. y Carter, C.S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, 8, 539-546.
- Bremmer, F., Schlack, A., Shah, N.J., Zafiris, O., Kubischik, M., Hoffmann, K.P., Zilles, K., Fink, G.R. (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron*, 29(1), 287-296.
- Brunet, D., Murray, M.M. y Michel, C.M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience*.
- Bushara, K.O., Grafman, J. y Hallett, M. (2001). Neural Correlates of Auditory-Visual Stimulus Onset Asynchrony Detection. *The Journal of Neuroscience*, 21(1), 300-304.
- Busse, L., Roberts, K.C., Crist, R.E., Weissman, D.H. y Woldorff, M.G. (2005). The spread of attention across modalities and space in a multisensory object *Proceedings of the National Academy of Sciences*, 102(51), 18751-18756.
- Caclin, A., Soto-Faraco, S., Kingstone, A. y Spence, C. (2002). Tactile "capture" of audition. *Perception & Psychophysics*, 64(4), 616-630.

- Calvert, G.A., Brammer, M.J., Bullmore, E.T., Campbell, R., Iversen, S.D. y David, A.S. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*, *10*, 2619-2623.
- Calvert, G.A., Campbell, R., y Brammer, M.J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*, 649-657.
- Calvert, G.A. (2001). Crossmodal Processing in the Human Brain: Insights from Functional Neuroimaging Studies. *Cerebral Cortex*, *11*, 1110-1123.
- Calvert, G.A., Spence, C. y Stein, B.E. (Eds.) (2004). *The handbook of multisensory processes*. Cambridge, MA: MIT Press.
- Calvert, G.A. y Thesen T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology*, *98*, 191-205.
- Cappe, C. y Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience*, *22*, 2886-2902.
- Cappe, C., Thut, G., Romei, V. y Murray, M.M. (2010). Auditory-visual multisensory interactions in humans: timing, topography, directionality, and sources. *Journal of Neuroscience*, *30*(38), 12572-12580.
- Cappe, C., Thelen, A., Romei, V., Thut, G. y Murray, M.M. (2012). Looming signals reveal synergistic principle of multisensory integration. *Journal of Neuroscience*, *32*, 1171-1182.
- Chen, K.M. y Yeh, S.L. (2009). Asymmetric cross-modal effects in time perception. *Acta Psychologica*, *130*, 225-234.
- Choe, C.S., Welch, R.B., Gilpord, R.M., y Juola, J.F. (1975). The "ventriloquist effect": Visual dominance or response bias? *Perception & Psychophysics*, *18*, 55-60.
- Colavita, F.B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*, 409-412.
- Corballis, P.M. y Gratton, G. (2003). Independent control of processing strategies for different locations in the visual field. *Biological Psychology*, *64*, 191-209.
- Corbetta, M. Patel, G. y Shulman, G.L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron*, *58*(3), 306-24.
- Correa, A., Lupiañez, J. y Tudela, P. (2006). La percepción del tiempo: Una revisión desde la Ciencia Cognitiva. *Cognitiva*, *18*, 145-168.

REFERENCIAS BIBLIOGRÁFICAS

- Crump, M.J., Gong, Z. y Milliken, B. (2006). The context-specific proportion congruent stroop effect: location as a contextual cue. *Psychonomic Bulletin & Review*, *13*, 316-321.
- Crump, M.J.C, Vaquero, J.M.M. y Milliken, B. (2008). Context-specific learning and control: The role of awareness, task set, and relative salience. *Consciousness & Cognition*, *17*, 22-36.
- Crump, M.J.C y Milliken, B. (2009). The flexibility of context-specific control: Evidence for context-driven generalization of item-specific control settings. *The Quarterly Journal of Experimental Psychology*, *62*, 1523-1532.
- Cusick, C.G. (1997). The superior temporal polysensory region in monkeys. En K.S. Rockland y J.H. Kaas (Eds.). *Cerebral cortex: extrastriate cortex in primates*. New York: Plenum.
- Davidson, D.J., Zacks, R.T. y Williams, C.C. (2003) Stroop interference, practice, and aging. *Aging, Neuropsychology and Cognition*, *10*, 85-98.
- de Jong, R., Toffanin, P., y Harbers, M. Dynamic crossmodal links revealed by steady-state responses in auditory-visual divided attention. *International Journal of Psychophysiology*, *75*, 3-15.
- De Lucia, M., Clarke, S. y Murray, M.M., (2010a). A temporal hierarchy for conspecific vocalization discrimination in humans. *Journal of Neuroscience*, *30* (33), 11210-11221.
- De Lucia, M., Michel, C.M. y Murray, M.M., (2010b). Comparing ICA-based and single-trial topographic ERP analyses. *Brain Topography*, *23* (2), 119-127.
- Deneve, S. y Pouget, A. (2004). Bayesian multisensory and cross-modal spatial links. *Journal of Physiology-Paris*, *98*(1-3), 249-58.
- Desimone, R. y Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*, 193-222.
- Dhamala, M., Assisi, C.G., Jirsa, V.K., Steinberg, F.L. y Kelso, J.A.S. (2007). Multisensory intergation for timing engages different brain networks. *Neuroimage*, *34*, 764-773.
- Diaconescu, A.O., Hasher, L. y McIntosh, A.R. (2013). Visual dominance and multisensory integration changes with age. *Neuroimage*, *15*(65), 152-166.
- Driver, J., y Spence, C. J. (1994). Spatial synergies between auditory and visual attention. En C. Umiltà y M. Moscovitch (Eds.). *Attention and performance XV: Conscious and nonconscious information processing*. Cambridge, MA: MIT Press.
- Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature*, *381* (6577), 66-68.

- Driver, J., y Spence, C. (1998). Crossmodal attention. *Current Opinion in Neurobiology*, 8, 245-253
- Duncan, J., Martens, S. y Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, 387(6635), 808-10.
- Duncan, J. (2013). The structure of cognition: attentional episodes in mind and brain. *Neuron*, 80(1), 35-50.
- Eagleman, D.M., Tse, P.U., Buonomano, D.V., Janssen, P., Nobre, A.C. y Holcombe, A. O. (2005). Time and the brain: How subjective time relates to neural time. *The Journal of Neuroscience*, 25, 10369-10371.
- Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, 12, 374-380.
- Egner, T. y Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, 8, 1784-1790.
- Ernst, M.O., Banks, M.S. y Bühlhoff, H.H. (2000). Touch can change visual slant perception. *Nature Neuroscience*, 3(1), 69-73.
- Ernst, M.O. y Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429-433.
- Ernst, M. O., & Bühlhoff, H. H. (2004). Merging the Senses into a Robust Percept. *Trends in Cognitive Sciences*, 8 (4), 162-169.
- Fairhall, S. L., y Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. *European Journal of Neuroscience*, 29, 1247-1257.
- Falchier, A., Clavagnier, S., Barone, P., y Kennedy, H. (2003). Anatomical evidence of multimodal integration in primate striate cortex. *Journal of Neuroscience*, 22, 5749-5759.
- Fort, A., Delpuech, C., Pernier, J. y Giard M.H. (2002). Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. *Cerebral Cortex*, 12 (10), 1031-1039.
- Johnson, J.A. y Zatorre, R.J. (2005). Attention to Simultaneous Unrelated Auditory and Visual Events: Behavioral and Neural Correlates. *Cerebral Cortex*, 15, 1609-1620.
- Ghazanfar, A.A. y Schroeder, C.E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, 10, 278-285.
- Getzmann, S. (2007). The effect of brief auditory stimuli on visual apparent motion. *Perception*, 36, 1089-1103.

REFERENCIAS BIBLIOGRÁFICAS

- Giard M.H. y Peronnet F. (1999). Auditory visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11, 473-490.
- Graziano, M.S.A, Reiss, L.A. y Gross, C.G. (1999). A neuronal representation of the location of nearby sounds. *Nature*, 397, 428-430.
- Grave de Peralta Menendez, R., Gonzalez Andino, S., Lantz, G., Michel, C.M., y Landis, T. (2001). Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topography*, 14 (2), 131-137.
- Grave de Peralta, R. y Gonzalez Andino S. (2002). Comparison of algorithms for the localization of focal sources: Evaluation with simulated data and analysis of experimental data. *International Journal of Bioelectromagnetism*, 4 (1), online journal.
- Grave de Peralta Menendez, R., Murray, M.M., Michel, C.M., Martuzzi, R., y Gonzalez Andino, S.L. (2004). Electrical neuroimaging based on biophysical constraints. *Neuroimage*, 21 (2), 527-539.
- Green, D.M., y Swets, J.A. (1966). Signal detection theory and psychophysics. New York: Wiley.
- Guthrie, D. y Buchwald, J.S. (1991). Significance testing of difference potentials. *Psychophysiology*, 28 (2), 240-244.
- Hackett, T.A., de la Mothe, L.A., Ulbert, I., Karmos, G., Smiley, J.F. y Schroeder, C.E. (2007). Multisensory convergence in auditory cortex, II. Thalamocortical connections of the caudal superior temporal plane. *The Journal of Comparative Neurology*, 502, 924-952.
- Hadjikhani, N. y Roland, P.E. (1998). Cross-modal transfer of information between the tactile and the visual representations in the human brain: a positron emission tomographic study. *Journal of Neuroscience*, 18, 1072-1084.
- Heinemann, A., Kunde, W. y Kiesel, A. (2009). Context-specific prime-congruency effects: On the role of conscious stimulus representations for cognitive control. *Consciousness and Cognition*, 18, 966-976.
- Ho, C., Reed, N. y Spence, C. (2007). Multisensory in-car warning signals for collision avoidance. *Human Factors*, 49, 1107-1114.
- Ho, C., & Spence, C. (2008). *The multisensory driver: Implications for ergonomic car interface design*. Aldershot: Ashgate Publishing.
- Holmes, N.P. (2009). The Principle of Inverse Effectiveness in Multisensory Integration: Some Statistical Considerations. *Brain Topography*, 21, 168-176.

- Jacoby, L. L., Lindsay, D.S., y Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review*, *10*, 634-644
- Jiménez, L. y Méndez, C. (1999). Which attention is needed for implicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 236-259.
- Kayser, C., Petkov, C.I. y Logothetis, N.K. (2009). Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. *Hearing Research*, *258* (1-2), 80-88.
- Kersten, D., Mamassian, P. y Yuille, A. (2004). Object perception as Bayesian inference, *Annual Review of Psychology*, *55*, 271-304.
- Kinchla, R.A. (1974). Detecting target elements in multielement arrays: A confusability model. *Perception & Psychophysics*, *15*(1), 149-158.
- King, J.A., Korb, F.M., y Egner, T. (2012). Priming of control: implicit contextual cuing of top-down attentional set. *Journal of Neuroscience*, *32*(24), 8192-8200.
- Klink, P.C., Montijn, J. S. y van Wezel, A. (2011). Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates. *Attention, Perception & Psychophysics*, *73*, 219-236.
- Koelewijn, T., Bronkhorst, A., y Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: a review of audiovisual studies. *Acta Psychologica*, *134*(3), 372-384.
- Koenig, T. y Melie-Garcia, L. (2010). A method to determine the presence of averaged event-related fields using randomization tests. *Brain Topography*, *23* (3), 233-242.
- Koppen, C. y Spence, C. (2007a). Audiovisual asynchrony modulates the Colavita visual dominance effect. *Brain Research*, *1186*, 224-232.
- Koppen, C. y Spence, C. (2007b). Spatial coincidence modulates the Colavita visual dominance effect. *Neuroscience Letters*, *417*, 107-111.
- Làdavas, E., Di Pellegrino, G., Farnè, A. y Zeloni, G. (1999). Neuropsychological evidence of an integrated visuotactile representation of peripersonal space in humans. *Journal of Cognitive Neuroscience*, *10*(5), 581-589.
- Larsen, A. McIlhagga, W. Baert, J. y Bundesen, C. (2003). Seeing or hearing? Perceptual independence, modality confusions, and crossmodal congruity effects with focused and divided attention. *Perception & Psychophysics*, *65*(4), 568-574.
- Lehmann, D. y Skrandies, W. (1980). Reference-free identification of components of checkerboard evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology*, *48*, 609-621.

REFERENCIAS BIBLIOGRÁFICAS

- Lehmann, D. (1987). Principles of spatial analysis. En Gevins, A.S. y Remond, A. (Eds.). *Handbook of electroencephalography and Clinical Neurophysiology: Methods of Analysis of Brain Electrical and Magnetic Signals, vol. 1*. Amsterdam: Elsevier.
- Lewis, J.W. y Van Essen, D.C. (2000). Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *The Journal of Comparative Neurology*, 428 (1), 112-137.
- Lindsay, D.S. y Jacoby, L.L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception & Performance*, 20, 219-234.
- Logan, G.D. y Zbrodoff, N.J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, 7, 166-174.
- Luck, S.J. (2005). An Introduction to the Event-Related Potential Technique. Cambridge, MA: MIT Press
- Luppino, G., Murata, A., Govoni, P. y Matelli, M. (1999). Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Experimental Brain Research*, 128, 181-187.
- Macaluso, E., George, N., Dolan, R., Spence, C. y Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET Study. *Neuroimage*, 21(2), 725-732.
- Macaluso, E. y Driver, J. (2005). Multisensory spatial interactions: a window onto functional integration in the human brain. *TRENDS in Neurosciences*, 28(5), 264-271.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.
- Marks, L.E. (2004). Cross-modal interactions in speeded classification. En G. A. Calvert, C. Spence. y B.E. Stein (Eds.). *Handbook of multisensory processes*. Cambridge, MA: MIT Press
- May, P.J. (2006). The mammalian superior colliculus: laminar structure and connections. *Progress in Brain Research*, 151, 321-378.
- McGurk, H. y MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746-748.
- McHaffie, J.G., Fuentes-Santamaria, V., Alvarado, J.C., Fuentes Farías, A.L., Gutiérrez-Ospina, G. y Stein, B.E. (2012). Anatomical features of the intrinsic circuitry underlying multisensory integration in the superior colliculus. En B.E. Stein (Ed.) *The new handbook of Multisensory Processes*. Cambridge (MA): MIT Press.

- Meredith M.A. y Stein B.E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, 221, 389-391
- Meredith M.A. y Stein B.E. (1983). Visual, Auditory, and Somatosensory Convergence on Cells in Superior Colliculus Results in Multisensory Integration. *Journal of Neurophysiology*, 56(3), 640-662.
- Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L. y Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, 115, 2195-2222.
- Michel, C.M., Koenig, T., Brandeis, D., Gianotti, L.R. y Wackermann, J. (2009). *Electrical neuroimaging*. Cambridge UK: Cambridge UP.
- Michel, C. M. y Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *Neuroimage*, 61, 371-385.
- Mishra, J., Martinez, A., Sejnowski, T.J. y Hillyard S.A. (2007). Early Cross-Modal Interactions in Auditory and Visual Cortex Underlie a Sound-Induced Visual Illusion. *The Journal of Neuroscience*, 27(15), 4120-4131.
- Mishra, J., Martínez A. y Hillyard S.A. (2010). Effect of Attention on Early Cortical Processes Associated with the Sound-Induced Extra Flash Illusion. *Journal of Cognitive Neuroscience*, 22, 1714-1729.
- Mishra, J. y Gazzaley, A. (2012). Attention Distributed across Sensory Modalities Enhances Perceptual Performance. *The Journal of Neuroscience*, 32(35), 12294 - 12302.
- Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E y Foxe J.J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognitive Brain Research*, 14(1), 115-128.
- Morein-Zamir, S., Soto-Faraco, S. y Kinstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cognitive Brain Research*, 17, 154-163.
- Murray, M.M., Michel, C.M., Grave de Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S. y Schnider, A. (2004). Rapid Discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage*, 21, 125-135.
- Murray, M.M., Brunet, D. y Michel, C.M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain Topography* 20, 249-264.
- Murray, M.M. y Wallace, M.T. (Eds) (2012). *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press.

REFERENCIAS BIBLIOGRÁFICAS

- Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W. y y Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Cognitive Brain Research*, 25(2), 499-507.
- Nickerson, R.S. (1973). Intersensory Facilitation of Reaction Time: Energy Summation or Preparation Enhancement. *Psychological Review*, 80 (6), 489-509.
- Padberg, J., Seltzer, B., y Cusick, C.G. (2003). Architectonics and cortical connections of the upper bank of the superior temporal sulcus in the rhesus monkey: an analysis in the tangential plane. *The Journal of Comparative Neurology*, 467, 418-434.
- Pascual-Leone, A., y Hamilton, R. (2001). The metamodal organization of the brain. *Progress in Brain Research*, 134, 427-445.
- Perrault, T.J., Vaughan, J.W., Stein, B.E. y Wallace, M.T. (2003). Neuron-specific response characteristics predict the magnitude of multisensory integration. *Journal of Neurophysiology*, 90, 4022-4026.
- Perrin, F., Pernier, J., Bertrand, O. y Echallier, J.F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and clinical Neurophysiology*, 72(2), 184-187.
- Petersen, S.E. y Posner, M.I. (2012). The Attention System of the Human Brain: 20 Years After. *Annual Review of Neuroscience*, 35, 73-89.
- Pick, H.L., Warren, D.H., y Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Perception & Psychophysics*, 6, 203-205.
- Pillai, A.S., Gilbert, J.R. y Horwitz, B. (2013). Early sensory cortex is activated in the absence of explicit input during crossmodal item retrieval: Evidence from MEG. *Behavioural Brain Research*, 238, 265-272
- Posner, M.I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, 32(1), 3-25.
- Pouget, A., Deneve, S. y Duhamel, J.R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience*, 3, 741-747.
- Pourtois, G., Delplanque, S., Michel, C. y Vuilleumier, P. (2008). Beyond conventional event-related brain potential (ERP): exploring the time-course of visual emotion processing using topographic and principal component analyses. *Brain Topography*, 20 (4), 265-277.
- Raij, T., Uutela, K y Hari, R. (2000). Audiovisual Integration of Letters in the Human Brain. *Neuron*, 28, 617-625.

- Rockland K.S., Ojima H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology*, 50, 19-26.
- Romei, V., De Haas, B., Mok, R.M. y Driver, J. (2011). Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. (2011). *Frontiers in Psychology: Perception Science*, 2(215), 1-8.
- Ross, L.A., Saint-Amour, D., Leavitt, V.N., Javitt, D.C. y Foxe, J.J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex*, 17, 1147-1153.
- Rushworth, M.F., Paus, T. y Sipila, P. K. (2001). Attention systems and the organization of the human parietal cortex. *Journal of Neuroscience*, 21, 5262-5271.
- Rugg, M.D. y Coles, M.G.H. (1996). *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*. Oxford, UK: Oxford University Press.
- Sanabria, D., Soto-Faraco, S. y Spence, C. (2007a). Spatial attention and audiovisual interactions in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, 33 (4), 927-937.
- Sanabria, D., Spence, C. y Soto-Faraco, S. (2007b). Perceptual and decisional contributions to audiovisual interactions in the perception of apparent motion: A signal detection study. *Cognition*, 102, 299-310.
- Sarmiento, B.R., Shore, D., Milliken, B., y Sanabria, D. (2012). Audiovisual interactions depend on context of congruency. *Attention Perception & Psychophysics*, 74, 563-574.
- Schmahmann, J.D., Pandya, D.N. (1991). Projections to the basis pontis from the superior temporal sulcus and superior temporal region in the rhesus monkey. *The Journal of Comparative Neurology*, 308, 224-248.
- Senkowski, D., Talsma, D., Herrmann, C.S. y Woldorff, M.G. (2005). Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. *Experimental Brain Research*, 166, 411-426.
- Senkowski, D., Talsma, D., Grigutsch, M., Herrmann, C.S. y Woldorff, M.G. (2007). Good times for multisensory integration: Effects of the precision of temporal synchrony as revealed by gamma-band oscillations. *Neuropsychologia*, 45, 561-571.
- Sekuler, R., Sekuler, A. B., y Lau, R. (1997). Sound alters visual motion perception. *Nature* 385, 308.
- Shams, L. Kamitani, Y. y Shimojo, S. (2000). What you see is what you hear. *Nature*, 408, 788.

REFERENCIAS BIBLIOGRÁFICAS

- Shams, L. Kamitani, Y. y Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, 14, 147-152.
- Shams L. y Kim R. (2010). Crossmodal influences on visual perception. *Physics of Life Reviews* 7, 269-284.
- Shimojo, S., y Shams, L. (2001). Sensory modalities are not separate modalities: Plasticity and interactions. *Current Opinion in Neurobiology*, 11, 505-509.
- Shore, D.I. y Simic, N. (2005). Integration of visual and tactile stimuli: top-down influences require time. *Experimental Brain Research*, 166 (3-4), 509-517.
- Sinnett, S., Soto-Faraco, S. y Spence, C. (2007). The co-occurrence of multisensory competition and facilitation. *Acta psychologica*, 128(1), 153-161.
- Slutsky, D.A., y Recanzone, G.H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*, 12, 7-10.
- Schmitt, M., Postma, A., y de Haan, E. (2000). Interactions between exogenous auditory and visual spatial attention. *Quarterly Journal of Experimental Psychology*, 53A, 105-130.
- Schmitt, M., Postma, A. y de Haan, E. (2001). Cross-modal exogenous attention and distance effects in vision and hearing. *European Journal of Cognitive Psychology*, 13, 343-368.
- Soto-Faraco, S., Spence, C., y Kingstone, A. (2004). Cross-Modal Dynamic Capture: Congruency Effects in the Perception of Motion Across Sensory Modalities. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 330-345.
- Spape M. M. y Hommel B. (2008). He said, she said: episodic retrieval induces conflict adaptation in an auditory Stroop task. *Psychonomic Bulletin & Review*, 15(6), 1117-1121.
- Spence C, Driver J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception & Performance*, 22(4), 1005-1030.
- Spence, C., Shore, D.I., y Klein, R.M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, 130(4), 799-832.
- Spence, C., y Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention*. Oxford, UK: Oxford University Press.
- Spence, C. y Ho, C. (2008). Multisensory driver interface design: Past, present, and future. *Ergonomics*, 51(1), 65-70.

- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception & Psychophysics*, 73, 971-995.
- Schröger, E., y Eimer, M. (1997). Covert spatial orienting in audition: 'Cost-Benefit' analyses of reaction times and event-related potentials. *The Quarterly Journal of Experimental Psychology*, 50A, 457-474.
- Stanford, T.R., Quessy, S. y Stein, B.E. (2005). Evaluating the Operations Underlying Multisensory Integration in the Cat Superior Colliculus. *Journal of Neuroscience*, 25(28), 6499-6508.
- Stein, B.E. y Meredith, M.A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.
- Stein, B. E., London, N., Wilkinson, L. K., & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *Journal of Cognitive Neuroscience*, 8, 497-506.
- Stein, B.E. (1998). Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Experimental Brain Research*, 123 (1-2), 124-135.
- Stein, B.E., Burr, D., Constantinidis, C., Laurienti, P.J., Meredith, M.A., Perrault Jr, T.J., Ramachandran, R., Röder, B., Rowland, B.A., Sathian, K., Schroeder, C.E., Shams, L., Stanford, T.R., Wallace, M.T., Yu. L. y Lewkowicz, D.J. (2010). Semantic confusion regarding the development of multisensory integration: a practical solution. *European Journal of Neuroscience*, 31, 1713-1720.
- Stein B.E, (Ed.) (2012). *The new handbook of multisensory processing*. Cambridge (MA): MIT Press.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.
- Stevenson, R.A., Ghose, D., Fister, J.K., Sarko, D.K., Altieri, N.A., Nidiffer, A.R., Kurela, L.R., Siemann, J.K., James, T.W. y Wallace, M.T. (2014). Identifying and Quantifying Multisensory Integration: A Tutorial Review. *Brain Topography*. Published online 11 April 2014.
- Sumby, W.H., y Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of Acoustic Society of America*, 26, 212-215.
- Talairach, J., y Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging. G. Thieme Medical Publishers, Stuttgart; New York.

REFERENCIAS BIBLIOGRÁFICAS

- Talati, A. y Hirsch, J. (2005). Functional specialization within the medial frontal gyrus for perceptual go/no-go decisions based on “what,” “when,” and “where” related information: an fMRI study. *Journal of Cognitive Neuroscience*, 17, 981-993.
- Talsma, D. y Woldorff, M.G. (2005). Selective attention and multisensory integration: multiple phases of effects on the evoked brain activity. *Journal of Cognitive Neuroscience*, 17 (7), 1098-1114.
- Talsma, D., Doty, T. J. y Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cerebral Cortex*, 17, 679-690.
- Talsma, D., Senkowski, D., Soto-Faraco, S. y Woldorff, M.G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14 (9), 400-410.
- Treisman, A, y Gelade, G. (1980). A Feature Integration Theory of Attention. *Cognitive Psychology*, 12, 97-136.
- Treisman, A, y Schmidt, H. (1982). Illusory Conjunctions in the Perception of Objects. *Cognitive Psychology*, 14, 107-141.
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6, 171-178.
- Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 353(1373), 1295-1306.
- Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition*, 14 (4-8), 411-443.
- Thelen, A., Cappe, C. y Murray, M. M. (2012). Electrical neuroimaging of memory discrimination based on single-trial multisensory learning, *NeuroImage* 62, 1478-1488.
- van Atteveldt, N., Murray, M.M., Thut, G., y Schroeder, C.E. (2014). Multisensory integration: flexible use of general operations. *Neuron*, 81, 1240-1253.
- Van der Burg, E., Olivers, C.N.L, Bronkhorst, A.W. y Theeuwes, J. (2008). Pip and Pop: Nonspatial Auditory Signals Improve Spatial Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, 34(5), 1053-1065.
- Vatakis, A. y Spence, C. (2007). Crossmodal binding: Evaluating the ‘unity assumption’ using audiovisual speech stimuli. *Perception & Psychophysics*, 69, 744-756.
- Vaughan, H.G.J., 1982. The neural origins of human event-related potentials. *Annals of the New York Academy of Sciences*, 388, 125-138.

- Verguts, T. y Notebaert, W. (2009). Adaptation by binding: a learning account of cognitive control. *Trends in Cognitive Sciences*, 13(6), 252-257.
- Vroomen, J. y de Gelder, B. (2000). Sound Enhances Visual Perception: Cross-modal Effects of Auditory Organization on Vision. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1583-1590.
- Vroomen, J., y Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics*, 72, 871-884.
- Wallace, M.T., Meredith, M.A. y Stein, B.E. (1993). Converging influences from visual, auditory y somatosensory cortices onto output neurons of the superior colliculus. *Journal of Neurophysiology*, 69 (6), 1797-1809.
- Walker, J.T. y Scott, K.J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones, and gaps. *Journal of Experimental Psychology: Human, perception and performance*, 7, 1327-1339.
- Watkins, S., Shams, L., Josephs, O. and Rees, G. (2007). Activity in human V1 follows multisensory perception. *Neuroimage*, 37, 572-578.
- Warren, D.H., Welch, R.B., y McCarthy, T.J. (1981). The role of visual-auditory “compellingness” in the ventriloquism effect: implications for transitivity among the spatial senses. *Perception & Psychophysics*, 30, 557-564.
- Welch, R.B., y Warren, D.H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, 88, 638-667.
- Welch, R. B. (1999). Meaning, attention, and the “unity assumption” in the intersensory bias of spatial and temporal perceptions. En G. Aschersleben, T. Bachmann, y J. Müsseler (Eds.). *Cognitive contributions to the perception of spatial and temporal events*. Amsterdam: Elsevier.
- Witten, I.B. y Knudsen, E.I. (2005). Why seeing is believing: merging auditory and visual worlds. *Neuron*, 48, 489-496.
- Zangenehpour, S. y Zatorre, R.J. (2010). Crossmodal recruitment of primary visual cortex following brief exposure to bimodal audiovisual stimuli. *Neuropsychologia*, 48(2), 591-600.
- Zatorre, R.J. y Belin, P. (2001). Spectral and Temporal Processing in Human Auditory Cortex. *Cerebral Cortex*, 11 (10), 946-953.

ANEXO I. EXPERIMENT 5

Method and results of Experiment 5 are described below. This study constituted the first experiment of Experimental Serie II, which was not included in the main text of this thesis due to technical issues with the EEG system during registration (see below). However, the report of Experiment 5 helps to understand some of the changes in the methods from Experiment 4 to Experiment 6, as it has been detailed in the “Motivación y objetivos” section.

1. Experiment 5

1.1. Method

1.1.1. Participants

The participants were 37 undergraduate students from the University of Granada (22 females; age range: 19-39 years old; mean age of 23 years old) who received course credits or cash payment of 10 Euro in exchange of their participation. All of the participants reported normal hearing and normal or corrected-to-normal vision, and gave their informed consent to participate in the study conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki.

1.1.2. Apparatus, stimuli and procedure

These were the same as in Experiment 6 except for the following: 1) the oddball task was not included in Experiment 5, therefore participants performed the visual discrimination task right after the unisensory task; 2) auditory and visual stimuli could last for 40 or 120 ms (see Figura 18); 3) participants completed 8 blocks of 80 multisensory trials each, distributed as follows: 32 congruent trials and 8 incongruent trials for the high proportion-congruent context; 8 congruent and 32 incongruent trials for the low proportion-congruent context; 4) Continuous EEG were acquired at 250 Hz.

EXPERIMENT 5

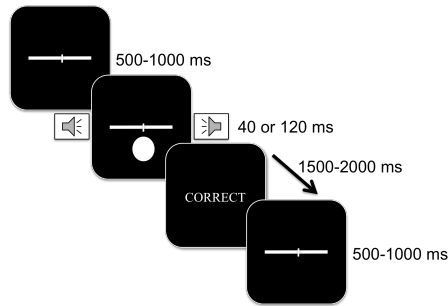


Figure 18. Schematic illustration of the set up used in the Experiment.

1.2. Results

The data from four participants were removed from the analysis because their performance did not reach 60% accuracy in the baseline condition. Additionally, eight participants did not reach a minimum of 34 trials per condition due to EEG artifacts, and the data were also removed. The data from the remaining 25 participants (15 women; age range: 20-39 years old, mean age 23) were included in the complete data analysis.

2.1.1. Behavioural data

Mean response accuracy for each participant and condition were submitted to a 2 Context (high congruent, low congruent) x 2 Congruence (congruent, incongruent) ANOVA, which revealed a significant main effect of Congruence, $F(1,24) = 41.13$, $MSE = .93$, $p < .001$, $\eta_p^2 = .63$. Responses on congruent trials (89% correct) were more accurate than responses on incongruent trials (69% correct). Both the main effect of Context and the interaction between Context and Congruence failed to reach statistical significance, both $F_s < 1$.

To investigate whether context specific proportion-congruent effect relied on trial-by-trial shifts in control setting, we performed a 2 Context transition (repetition, switch) x 2 Context (high congruent, low congruent) x 2 Congruence (congruent, incongruent)

ANOVA³ (see King et al., 2012, for a similar approach). The first trial of each block was excluded from the analysis. Crucially, this further ANOVA showed a significant three-way interaction, $F(1,24) = 21.35$, $MSE = .093$, $p < .001$, $\eta_p^2 = .47$ (See Figure 19). To examine differences in performance as a function of Context transition, a 2 Context (high congruent, low congruent) by 2 Congruence (congruent, incongruent) ANOVA was conducted for each type of transition, separately. Results revealed a significant two-way interaction for both conditions, whenever stimulus presentation followed a context repetition $F(1,24) = 14.31$, $MSE = .06$, $p < .01$, $\eta_p^2 = .37$ and when stimulus display switched between contexts $F(1,24) = 13.89$, $MSE = .035$, $p = .001$, $\eta_p^2 = .37$. Further analysis showed a reliable context specific proportion-congruent effect on the former condition due to lower accuracy on congruent trials (85,69%), $F(1,24) = 13.64$, $MSE = .026$, $p < .01$, and higher accuracy on incongruent trials (72,03%), $F(1,24) = 9.04$, $MSE = .033$, $p < .01$, in the low proportion-congruent context with respect to the high proportion-congruent context (90,34% and 66,99 for congruent and incongruent respectively). The inverse pattern of results was observed in the condition where context of congruence switched between trials: an increased accuracy on incongruent trials (71,8%), $F(1,24) = 8.49$, $MSE = .031$, $p < .01$, and an accuracy decrement in the congruent trials (88,01%), $F(1,24) = 6.44$, $MSE = .008$, $p < .02$, in the high proportion-congruent context relative to the low proportion-congruent context (66,81% and 90,51% for incongruent and congruent respectively). In terms of the magnitude of the congruence effects, these results replicate previous accounts (e.g., King et al., 2011) showing the relevance not only of the current context of congruence but of the context of congruence of the previous trial.

³ This analysis could not be run on the electrophysiological data because once the experiment was split as a function of context transition the remaining trials on each condition were not enough to preserve the signal-to-noise ratio.

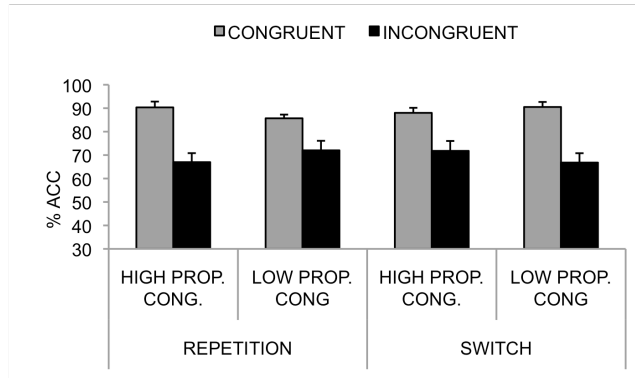


Figure 19. Percentage of correct responses as a function of Context transition, Context and Congruence. Error bars represent the standard error of the mean.

2.1.1. Questionnaire measures

Participant's reports revealed that 8 (32%) of the participants noticed a difference between the two contexts (e.g., "I made more errors on the upper half"). However none of them reported being aware of the specific proportion-congruent manipulation, in accordance to previous studies (e.g., Crump et al., 2006; Heinemann et al., 2009; Sarmiento et al., 2012; King et al., 2012).

2.1.2. Surface ERP data

A point-wise 2 Context (high congruent, low congruent) x 2 Congruence (congruent, incongruent) ANOVA on the Global Field Power data was conducted to determine global modulations in the strength of the electric field. This analysis provided evidence of a significant interaction between the two factors during two phases, a first period over ~116-216 ms and a second period covering the ~224-280 ms (see Figure 20). Post-hoc comparisons on the ~116-216 period revealed a stronger response on congruent trials than incongruent trials in the high proportion-congruent context, while a stronger response on incongruent trials than on congruent trials was shown for the low proportion-congruent context.

Over the ~224-280 ms interval, post-hoc comparisons showed stronger response on congruent trials than incongruent trials for the high proportion-congruent context, while

there was no statistical difference between congruent and incongruent trials in the low proportion-congruent context.

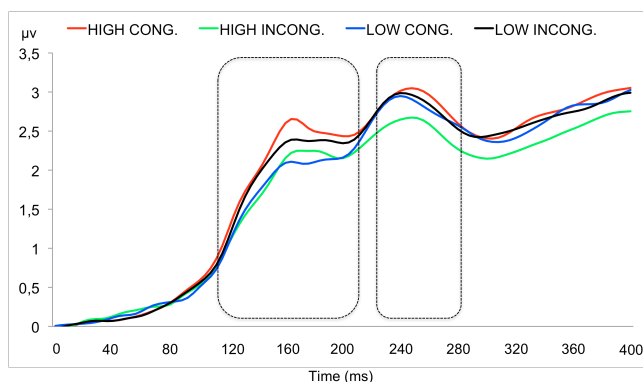
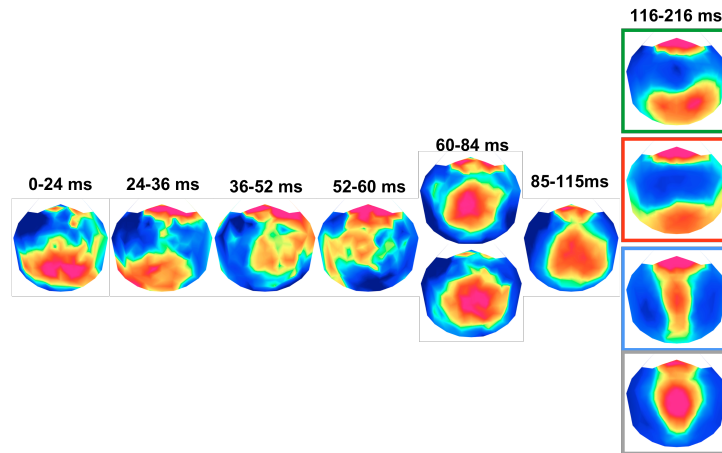


Figure 20. Group-average GFP waveforms and results of the ANOVA on the GFP as a function of time ($p < 0.05$; $> 12ms$) Dotted boxes represent intervals of statistical significant differences.

A hierarchical topographic cluster analysis was performed on the collective group-average ERPs to identify periods of stable electric field topography both within and between experimental conditions. 20 template maps that explained 98.34% of the variance were identified. At the group-average level two main results were observed (see Figure 21a). Different map configurations over the ~116-216 ms post-stimulus period accounted for each context of congruence, and a faster onset for the last map as a function of the context of congruence. This pattern was statistically verified in the single-subject ERPs using a spatial-correlation fitting procedure over the periods of interest. A 2 Context x 2 Congruence x 4 Map ANOVA revealed a significant interaction between Context and Map for the ~116-216 ms period, $F(3,72) = 3.49$, $MSE = 76.41$, $p < 0.02$, $\eta_p^2 = .13$. A priori comparisons showed that the green map predominated in responses in the low proportion-congruent context $F(1,24) = 5.02$, $MSE = 104.04$, $p < 0.04$, while the red map accounted for the responses in the high proportion-congruent context $F(1,24) = 4.51$, $MSE = 118.81$, $p < 0.05$. (see Figure 21b).

EXPERIMENT 5

a



b

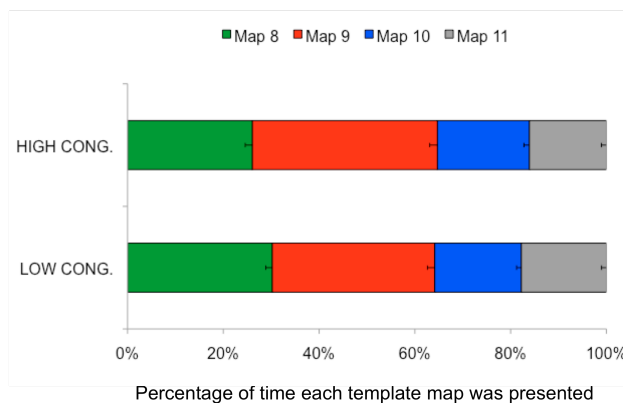


Figure 21. a Template maps from topographic cluster analysis. **b.** Results of the spatial correlation fitting in the single-subject data from each condition. Error bars represent the standard error of the mean

2.1.3. Source estimations

The source estimations from the ~116-216 ms time period were statistically analyzed to estimate the intracranial generators contributing to the differences between the low proportion-congruent context and the high proportion-congruent context conditions. Results revealed a main effect of Context (see Figure 22). Particularly, two clusters showed the strongest significant differences. Those clusters were located in the left

anterior superior temporal gyrus (Max: -47, 0, -8 mm; BA 21) and the left parahippocampal gyrus (Max: -23, -29, -11, BA36) extending into the hippocampus.

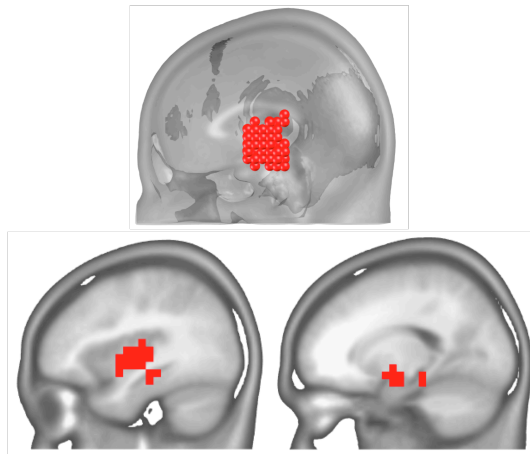


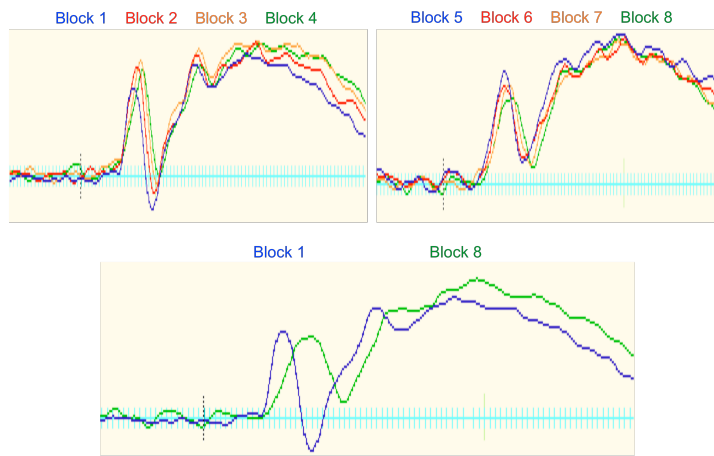
Figure 22. Results of source estimations ANOVA between the high and low proportion-congruent contexts over the 116-216 ms interval. Note that the positive values (warmer colors) indicate that the clusters are more active in the high proportion-congruent condition than in the low proportion-congruent condition.

2. Technical issues

The problem that occurred with Net Station software is still unknown, but it was manifested as a delay in the registration of the events throughout the different trials of the experiment. As can be seen in Figure 22a, the first peak of the waveform was displaced from one block to the next in Experiment 5. This delay was not present in Experiment 6 (see Figure 22b). Note that the explanation that such delay could be an effect or an artifact due to our manipulation is excluded, since it seems that this pattern of delay was present in more studies conducted in the lab during that period of time.

a

Experiment 5



b

Experiment 6



Figure 23. Waveforms of Experiment 5 (a) and Experiment 6 (b)