Voluntary modulations of attention in a semantic auditory-visual matching Task: an ERP study

RODRIGO ORTEGA^{1,2}, VLADIMIR LÓPEZ^{1,3} and FRANCISCO ABOITIZ¹

¹ Laboratorio de Neurociencia Cognitiva, Departamento de Psiquiatría, Centro de Investigaciones Médicas, Facultad de Medicina, Pontificia Universidad Católica de Chile.

² Escuela de Psicología, Universidad Diego Portales, Santiago, Chile.

³ Escuela de Psicología, Facultad de Ciencias Sociales, Pontificia Universidad Católica de Chile.

ABSTRACT

The present study explores the neural correlates of voluntary modulations of attention in an auditory-visual matching task. Visual stimuli (a female or a male face) were preceded in close temporal proximity by auditory stimuli consisting of the Spanish word for “man” and “woman” (“hombre” or “mujer”). In 80% of the trials the gender of the two stimuli coincided. Participants were asked to mentally count the specific instances in which a female face appeared after hearing the word “man” (10 % of the trials). Our results show attention-related amplitude modulation of the early visual ERP components N1 and anterior P2, but also amplitude modulations of (i) the N270 potential usually associated with conflict detection, (ii) a P300 wave related to infrequency, and (iii) an N400 potential related to semantic incongruence. The elicitation of these latter components varied according to task manipulations, evidencing the role of voluntary allocation of attention in fine-tuning cognitive processing, which includes basic processes like detection of infrequency or semantic incongruity often considered to be volition-independent.

key terms: a) Voluntary attention, b) semantic processing, c) Top down control, d) ERP components.

INTRODUCTION

Selective attention is known to exert a top-down modulation of cognitive mechanisms at different levels. The system described as the voluntary network of attention constitutes one of the main mechanisms for redistributing attentional resources according to a macrocontext or an experimental setting, for example determining the relevance of a specific stimulation condition in a task-solving oriented strategy (Kok, 2001; Corbetta and Shulman, 2002). In this context, voluntary attention has been shown to modify the timing and allocation of resources related to processes usually considered to be automatic (Yantis and Jonides, 1990). Attentional allocation is usually inferred from improved behavioral performance and

increased amplitude in attention-sensitive ERP components (Posner, 1980; Hillyard and Anllo-Vento, 1998).

Conflict detection, on the other hand, is a well characterized process that produces specific electrophysiological and hemodynamic responses that depend on stimulus complexity (van Veen and Carter, 2002; Mao and Wang, 2007). Recently, there has been increasing interest in conflict detection mechanisms during multimodal incongruity tasks (Larsen et al., 2003; Teder-Salejarvi et al., 2005). Although the relationship between attention and conflict detection has been explored in single modality tasks (Wang et al., 2003), their interplay in the case of crossmodal integration needs further research. Our goal in this study is to address the effects of voluntary attention in conflict detection

Corresponding Author: Francisco Aboitiz. Address: Laboratorio de Neurociencia Cognitiva, Centro de Investigaciones Médicas, Facultad de Medicina, Pontificia Universidad Católica de Chile, Marcoleta 391, Santiago, Chile. Phone: 562 354 3808, 562 638 0443. Fax: 562 665 1951. Email: faboitiz@uc.cl Web: www.neuro.cl

Received: August 7, 2008. In Revised form: November 18, 2008. Accepted: December 18, 2008

mechanisms, using a bimodal auditory-visual gender-matching task.

Event Related Potentials (ERP) has been widely used to study attention, crossmodal integration and conflict detection. Recently, Wang et al. (2002b) demonstrated the elicitation of a negative component around 270 ms related to incongruity (N270) in a visual-auditory intermodal integration task where male or female faces were contrasted with the voice of a male or a female pronouncing a vowel. The amplitude of this component increased when there was gender incongruence between the face and the voice. Along the same line, several studies report a family of negative ERP components associated with detection and processing of conflictive information, which are elicited depending on the number and nature of stimulus attributes in conflict (Wang et al., 2004). The N270 potential can be elicited in tasks where two visual stimuli differ in form (Cui et al., 2000), color (Tian et al., 2001), spatial localization (Yang and Wang, 2002) or other characteristics (Zhang et al., 2001; Wang et al., 2002a; Zhang et al., 2002; Zhang et al., 2003). Facing more complex conflicts, for example when stimuli differ in more than one attribute, the N270 wave is also elicited but followed by a later negative deflection peaking around 400 ms (Wang et al., 2004). This late negative component has been widely reported in association with semantic incongruence (Kutas and Hillyard, 1980a; Kutas and Federmeier, 2000). In this context, the absence of an N400 ERP component in the study by Wang et al. (2002b) described above, may reflect the fact that only phonological, and not semantic processing, of the auditory stimuli was necessary to solve the task.

The task relevance of the conflict has been also explored presenting relevant and irrelevant incongruent conditions (Wang et al., 2001). A larger N270 for the relevant condition has been reported and interpreted as evidence of an attentional enhancement of conflict processing activity when the conflict is relevant to the task, in spite of the usually automatic operation of this process (Wang et al., 2003). The present work was intended to address the following

questions: a) what is the role of voluntary allocation of attention in the conflict detection system? b) how does the relative salience of the conflict (relevant vs. irrelevant for the task) and its semantic nature (male vs. female) modify the activity of this system?

METHODS

Participants: Ten right-handed healthy volunteers, native Spanish speakers (five of them women), between 18 and 30 years of age, with normal or corrected-to-normal vision and no reported auditory problems were recruited. Informed consent was obtained from all participants.

Stimuli: Auditory and visual stimulation was accomplished using the STIM™ system (NeuroScan-Compumedics) synchronized with a digital electroencephalograph. Auditory stimuli (the words “man” and “woman” in Spanish language) were presented binaurally through insertion earphones at 90 dB spl. Visual stimuli consisted of images of a male or a female face presented at the center of a 21” computer screen (5° of visual field). The monitor was placed at a distance of 50 cm in front of the subject.

Procedure: Each trial began with a fixation cross presented in the center of the screen for 200 ms. This was followed by the auditory stimulus, which had an approximate duration of 500 ms. A silent and blank screen with a random duration of 50 – 150 ms separated the auditory stimulus from the visual presentation. Visual stimuli were presented for 250 ms and followed by a blank screen for 600 ms that preceded the next trial (Fig. 1).

The experimental design consisted of 240 trials. In 80% of the trials, auditory and visual input were coincident (either the word “man” was followed by a male face, or the word “woman” was followed by a female face), and in the other 20% of the trials, stimuli were incongruent (which was the word “man” followed by the female face, or the word “woman” was followed by a male face). Although there were two incongruent conditions, each appearing

10% of the time, participants were asked to pay attention and mentally count only the occurrence of the female face after hearing the word “man”. Thus, there were two equally incongruent and equally infrequent conditions, but one was being voluntarily attended (relevant condition) while the other was irrelevant to the task.

The task was conducted in two blocks of equal duration, separated by a resting period of at least 5 minutes. After each block, the participants reported the number of occurrences of the target conflictive pair,

which was the female face after the word “man”. All participants had accuracy reports above 90 % in both blocks. The experiment was conducted in afternoons for all subjects. They were asked to relax while they were prepared for the EEG recording session.

EEG Recording: Electrophysiological signals were recorded using a NeuroScan™ 80-channel digital electroencephalograph with high-resolution NuAmps™ amplifiers. An 80-channel cap (QuickCap™) from the same company was used for electrode

Fig. 1: Trial sequence and experimental conditions. Auditory stimulus was followed by a visual stimulus that could match or not regarding gender (Match – Mismatch conditions). Subjects were asked to pay attention and mentally count only the mismatch condition in which the word “man” was followed by a female face (task relevant –irrelevant conditions).

placement. Impedances were kept below 5 k Ω throughout the recordings. A/D sampling frequency was set at 250 Hz. A band pass digital filter between 0.5 and 30 Hz was later applied to remove unwanted frequency components. Two additional bipolar derivations were used to monitor vertical and horizontal ocular movements (EOG). Continuous EEG data were segmented between 200 ms prior to each visual stimulus and 800 ms after it. All segments with eye movement contamination, or any other technical or biological artifact, were removed from any further analysis. Artifact free segments were averaged to obtain the ERPs. The EEGLAB (Delorme and Makeig, 2004) Matlab toolbox was used for EEG off-line processing and analysis.

Statistical Analysis: All statistical calculations were performed using individual waveforms. A repeated measures ANOVA design with two factors (gender-match X task-relevance) was conducted, using the amplitude values of ERP components that differed between conditions. Each one of these factors had two levels (match-mismatch, relevant-irrelevant). Latency windows for statistic analysis of ERP effects were: N1 [170-200 ms], anterior P2 [170-200 ms], N270 [260-290 ms], P300 [360-400ms] and N400 [450-500 ms]. Results were corrected with Greenhouse-Geisser and Huynh-Feldt methods to adjust the univariate output of repeated measures ANOVA for violations of the compound symmetry assumption.

RESULTS

The ERPs recorded in the present experiment showed a series of negative and positive deflections that according to their peak latency, polarity, scalp topography and resemblance to classical effects reported in the literature were named as N1, P2, N270, P300 and N400 (Fig. 2). The amplitude of the N1 component was basically modulated by the photo-gender match or mismatch conditions, but the gender of the participants yielded no significant effect. This component exhibited larger amplitude

on posterior sites when elicited by a visual stimulus semantically congruent with the preceding auditory stimulus. This effect was less conspicuous when the same combination was presented in a task-relevant condition. The ANOVA showed a significant main effect for match $F_{(1,9)}=14.3$, $p<0.001$. Task relevance had marginal statistical significance $F_{(1,9)}=5.28$, $p=0.04$ but there were no significant interactions between these two factors $F_{(1,9)}=0.92$, $p=0.36$. In the same time window, but on anterior sites positivity becomes evident in the waveform corresponding to the task relevant and gender mismatch condition, i.e. the instructed target. This effect, named anterior P2, resulted statistically significant $F_{(2,8)}=15.3$, $p<0.001$.

A negative deflection was observed around 270 ms in both incongruent conditions, relevant and irrelevant. This N270 negativity had an increased amplitude in the gender mismatch condition in comparison to the match condition $F_{(2,8)}=16.9$, $p<0.001$. Task-relevance had no significant effect over N270 amplitude $F_{(2,8)}=0.74$, $p=0.50$ and there were no significant interactions $F_{(2,8)}=1.02$, $p=0.39$.

Additionally, two significant effects were found on the ERP waveforms corresponding to the task relevant and gender-mismatch condition. The first is positivity with a mean latency around 380 ms and central-frontal scalp distribution (Fig. 3). This P300 component had significantly larger amplitude in the task-relevant mismatch condition $F_{(3,7)}=31.02$, $p<0.001$. The second is a negative deflection with an approximate latency of 470 ms that was also elicited only in the task-relevant mismatch condition $F_{(3,7)}=15.26$, $p<0.001$ (see Table I).

DISCUSSION

The present results show attention related modulations at several levels of cortical processing, as indexed by the ERP components N1, P2, N270, P300 and N400. The cortical generators of these components have been extensively studied and while the

Fig. 2: Grand-average ERP waveforms depicted according to task relevance and match (solid line) or mismatch (dashed line) conditions. Midline electrodes Fcz, Cz and Oz. Amplitude scale in microvolts (-8 to 6 μ V), positive up. Time scale in milliseconds (-200 to 800 ms).

Fig. 3: Scalp topographic maps illustrating the amplitude distribution of the main ERP effects (N1-P2, N270, P300 and N400). Maps were calculated from the difference waveform mismatch – match in the task relevant condition. Amplitude scale in microvolts (μ V), positive up (lighter).

early components are generated in primary sensory cortices and nearby association areas (Di Russo et al, 2003), the N270 seems to be generated in the anterior cingulate (Li et al, 2003), and the late components (P300 and N400) have a more complex pattern of generation involving multiple areas (Kutas and Federmeier, 2000; Kok, 2001). The early N1 components exhibited larger amplitudes in conditions of match between visual and auditory information about gender. This effect was evident regardless of stimulus relevance, but in approximately the same time window a positive component became evident on anterior sites, and was only elicited by the relevant targets (task-relevant + gender mismatch). These effects are in line with previous results reported in single modality tasks that have been interpreted as enhancements due to an attentional set established prior to stimulus delivery (Potts et al., 2004). They also complement recent reports about early attentional modulation depending on the spatial location of auditory and visual stimuli (Nager et al., 2006; Schaefer et al., 2006). Nevertheless, our results show that this modulation occurs not only for spatially coincident stimuli but is also found when they concur in more complex features, including meaning.

The evidences presented here support the view that early attentional selection depends not only on the activity within a single sensory modality, but also on the activity from cortical areas belonging to other sensory systems and from multimodal integration areas, as has been recently argued by others (Talsma et al., 2007). Thus, early attentional selection can be influenced by previously established wider-frame cognitive sets like the expressed instructions to solve the task.

At later stages of processing, amplitude modulation of ERP components seems to reflect the involvement of other attention related processes. The task relevant condition elicited N270, P300 and N400 effects, but in the task irrelevant condition only the N270 was observed. The elicitation of the N270 by incongruent pairs of stimuli independently of their relevance to solve the task suggests that this component reflects the activity of a conflict detection process of automatic nature, rather resistant to top-down influences from voluntary attention networks. This result is in disagreement with a previous report by Wang et al. (Wang et al., 2001) that found that although the N270 was elicited independently of the task relevance, its amplitude increased when the conflicting attribute was relevant to the task. The

TABLE I

Mean and Standard Deviation (mean $\pm$ SD) values for each ERP component in each condition

Amplitude Values (μ V)		Relevant		Irrelevant	
Component	Channel	Match	Mismatch	Match	Mismatch
N1	Oz	-7.21** $\pm$ 0.78	-5.6 $\pm$ 0.62	-5.32** $\pm$ 0.56	-3.89 $\pm$ 0.87
P2	FCz	2.92 $\pm$ 0.16	4.43** $\pm$ 0.58	3.22 $\pm$ 0.43	3.52 $\pm$ 0.38
N270	FCz	-1.05 $\pm$ 0.51	-2.56** $\pm$ 0.66	-1.91 $\pm$ 0.85	-3.41** $\pm$ 0.53
	Cz	-2.04 $\pm$ 0.43	-3.03** $\pm$ 0.58	-0.63 $\pm$ 0.27	-2.97** $\pm$ 0.49
P300-like	FCz	0.82 $\pm$ 0.46	2.40** $\pm$ 0.33	-1.01 $\pm$ 0.52	-0.16 $\pm$ 0.59
	Cz	1.05 $\pm$ 0.32	2.72** $\pm$ 0.26	-0.93 $\pm$ 0.51	-0.11 $\pm$ 0.44
N400	FCz	-0.64 $\pm$ 0.53	-2.15** $\pm$ 0.49	-0.83 $\pm$ 0.28	-0.25 $\pm$ 0.36
	Cz	-0.39 $\pm$ 0.46	-1.58** $\pm$ 0.24	-1.15 $\pm$ 0.56	-0.27 $\pm$ 0.40

Two factors ("gender-match" x "task-relevance") repeated measures ANOVA, n=10. **P<0.001 (μ V: microvolts).

differences between the two experimental designs, especially the complexity of the task and the sensory modality, impede further comparisons regarding this point. The impact of the task complexity in the operation of this conflict detection system and its associated ERP effects have been documented before (Wang et al., 2004). It has also been argued that the N270 could be functionally similar to other ERP components like the N2b or N220 known to be elicited in tasks involving stimuli comparisons, but recent studies reaffirm its direct relation to conflict detection (Zhang et al., 2005; Szucs et al., 2007).

The presence of ERP effects like P300, usually related with stimulus infrequency (for a review see Kok, 2001), and N400, classically associated with semantic incongruence (Kutas and Hillyard, 1980b; Kutas y Federmeier, 2000) only in the incongruent-infrequent condition relevant to the task is in agreement with the high sensitivity to attentional factors previously reported for these components (Kutas and Federmeier, 2000; Kok, 2001). The salience of a stimulus due to other characteristics, for example frequency of presentation, has not been studied in this context before. Infrequency is known to capture and reorient attention and evokes the P300. All the previous reports that associate conflict detection with the elicitation of a sequence of negativities (N270 and N400) used the same frequency of presentation for coincident and non-coincident stimuli, perhaps in order to avoid confounding factors, but precluding the exploration of a highly common combination in real life: conflictive information also tends to be infrequent. The present design reaffirms the relevance of top-down attentional mechanisms in the elicitation or suppression of these components: the same type of incongruence and the same infrequency might or might not elicit N400 and P300, depending on a previously set "allocation policy" for cognitive resources. This pattern of complex resource allocation seems to be in complete accordance with a goal directed strategy.

Regarding the appearance of the sequence N270 and N400, our results

suggest that besides the complexity of conflict and the number of conflicting attributes, the need for semantic processing and the relevance of conflict detection to solve the task might be additional factors influencing the elicitation of the second negativity.

All these findings suggest that, aside from automatic conflict detection, voluntary attention is necessary for cognitive processing oriented to task resolution. Top-down modulations seem to modify the processing of stimulus related information not only at late stages but also at very early stages of attentional selection as reflected by the elicitation of the anterior P2. Further research is needed to clarify how these multiple processes come together in a coordinate manner, but the present work emphasizes the need of considering task instructions, a priori information and the task-solving strategy implemented by the participants as important elements to understand attention modulations in ERP experiments. Experience and available information play an important role tuning conflict detection mechanisms according to task demands.

ACKNOWLEDGEMENTS

Supported by Fondecyt Grants 1050721, 1080219 and by The Millenium Center for Integrative Neuroscience.

REFERENCES

- CORBETTA M, SHULMAN GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3: 201-215
- CUI L, WANG Y, WANG H, TIAN S, KONG J (2000) Human brain sub-systems for discrimination of visual shapes. *Neuroreport* 11: 2415-2418
- DELORME A, MAKEIG S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134: 9-21
- DI RUSSO F, MARTÍNEZ A, HILLYARD SA (2003) Source analysis of event-related cortical activity during visuo-spatial attention. *Cereb Cortex* 13: 486-499
- HILLYARD SA, ANLLO-VENTO L (1998) Event-related brain potentials in the study of visual selective attention. *Proc Natl Acad Sci U S A* 95: 781-787
- KOK A (2001) On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology* 38: 557-577

- KUTAS M, HILLYARD SA (1980a) Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biol Psychol* 11: 99-116
- KUTAS M, HILLYARD SA (1980b) Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207: 203-205
- KUTAS M, FEDERMEIER KD (2000) Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn Sci* 4: 463-470
- LARSEN A, MCILHAGGA W, BAERT J, BUNDESEN C (2003) Seeing or hearing? Perceptual independence, modality confusions, and crossmodal congruity effects with focused and divided attention. *Percept Psychophys* 65: 568-574
- LI L, WANG Y, WANG H, CUI L, TIAN S (2003) Event-related potential N270 and its distribution in adults and school-age children. *Brain Dev* 25: 507-513
- MAO W, WANG Y (2007) Various conflicts from ventral and dorsal streams are sequentially processed in a common system. *Exp Brain Res* 177: 113-121
- NAGER W, ESTORF K, MUNTE TF (2006) Crossmodal attention effects on brain responses to different stimulus classes. *BMC Neurosci* 7: 31
- POSNER MI (1980) Orienting of attention. *Q J Exp Psychol* 32: 3-25
- POTTS GF, PATEL SH, AZZAM PN (2004) Impact of instructed relevance on the visual ERP. *Int J Psychophysiol* 52: 197-209
- SCHAEFER M, TEWES U, MUNTE TF, JOHANNES S (2006) Lateralized irrelevant speech alters visuospatial selective attention mechanisms. *Biol Psychol* 72: 51-58
- SZUCS D, SOLTESZ F, CZIGLER I, CSEPE V (2007) Electroencephalography effects to semantic and non-semantic mismatch in properties of visually presented single-characters: the N2b and the N400. *Neurosci Lett* 412: 18-23
- TALSMA D, DOTY TJ, WOLDORFF MG (2007) Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cereb Cortex* 17: 679-690
- TEDER-SALEJARVI WA, DI RUSSO F, MCDONALD JJ, HILLYARD SA (2005) Effects of spatial congruity on audio-visual multimodal integration. *J Cogn Neurosci* 17: 1396-1409
- TIAN S, WANG Y, WANG H, CUI L (2001) Interstimulus interval effect on event-related potential N270 in a color matching task. *Clin Electroencephalogr* 32: 82-86
- VAN VEEN V, CARTER CS (2002) The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiol Behav* 77: 477-482
- WANG H, WANG Y, KONG J, CUI L, TIAN S (2001) Enhancement of conflict processing activity in human brain under task relevant condition. *Neurosci Lett* 298: 155-158
- WANG H, WANG Y, WANG D, CUI L, TIAN S, ZHANG Y (2002a) Cognitive impairment in Parkinson's disease revealed by event-related potential N270. *J Neurol Sci* 194: 49-53
- WANG Y, WANG H, CUI L, TIAN S, ZHANG Y (2002b) The N270 component of the event-related potential reflects supramodal conflict processing in humans. *Neurosci Lett* 332: 25-28
- WANG Y, TIAN S, WANG H, CUI L, ZHANG Y, ZHANG X (2003) Event-related potentials evoked by multi-feature conflict under different attentive conditions. *Exp Brain Res* 148: 451-457
- WANG Y, CUI L, WANG H, TIAN S, ZHANG X (2004) The sequential processing of visual feature conjunction mismatches in the human brain. *Psychophysiology* 41: 21-29
- YANG J, WANG Y (2002) Event-related potentials elicited by stimulus spatial discrepancy in humans. *Neurosci Lett* 326: 73-76
- YANTIS S, JONIDES J (1990) Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *J Exp Psychol Hum Percept Perform* 16: 121-134
- ZHANG X, WANG Y, LI S, HUANG X, CUI L (2002) Early detection of cognitive impairment in patients with obstructive sleep apnea syndrome: an event-related potential study. *Neurosci Lett* 325: 99-102
- ZHANG X, WANG Y, LI S, WANG L (2003) Event-related potential N270, a negative component to identification of conflicting information following memory retrieval. *Clin Neurophysiol* 114: 2461-2468
- ZHANG X, WANG Y, LI S, WANG L, TIAN S (2005) Distinctive conflict processes associated with different stimulus presentation patterns: an event-related potential study. *Exp Brain Res* 162: 503-508
- ZHANG Y, WANG Y, WANG H, CUI L, TIAN S, WANG D (2001) Different processes are involved in human brain for shape and face comparisons. *Neurosci Lett* 303: 157-160.