Cognitive entrainment to isochronous rhythms is independent of both sensory modality and top-down attention

Diana Cutanda, Daniel Sanabria & Ángel Correa[*]

Centro de Investigación Mente, Cerebro y Comportamiento,

Universidad de Granada (SPAIN)

The anisochrony of a stimulus sequence was manipulated parametrically to investigate whether rhythmic entrainment is stronger in the auditory modality than in the visual modality (Experiment 1), and whether it relies on top-down attention (Experiment 2). In Experiment 1, participants had to respond as quickly as possible to a target presented after a sequence of either visual or auditory stimuli. The anisochrony of this sequence was manipulated parametrically, rather than in an all or none fashion; that is, it could range from smaller to larger deviations of the isochrony (0, 10, 20, 50, 100, 150 and 200 ms). We compared rhythmic entrainment patterns for auditory and visual modalities. Results showed a peak of entrainment for both isochrony and deviations of isochrony up to 50 ms (i.e., participants were equally fast both after the isochronous sequences and after 10, 20 and 50 ms deviations), suggesting that anisochronous sequences can also produce entrainment. Beyond this entrainment window, the reaction times became progressively slower. Surprisingly, no differences were found between the entrainment patterns for auditory and visual rhythms. In Experiment 2, we used a dual-task methodology by adding a working memory n-back task to the procedure of Experiment 1. Results did not show interference of the secondary task in either auditory or visual modalities, with participants showing the same entrainment pattern as in Experiment 1. These results suggest that rhythmic entrainment constitutes a cognitive process that occurs by default (automatically), regardless of the modality in which the stimuli are presented, and independent of top-down attention, to generate behavioural benefits.

The environment provides us with many sources of temporal regularity to which we are exposed from an early age. Two and a half-year-old children can both synchronise their tapping to a regular tempo and anticipate beats adjusting their motor response if the given tempo becomes slower (Provasi & Bobin-Bègue, 2003). External regular rhythms can both facilitate motor synchronisation and provide information about when relevant events are going to occur. The generation of temporal expectancies is thus exogenously induced by regular patterns of stimuli (Nobre, Correa, & Coull, 2007). These temporal predictions result in behavioural benefits in response to the events matching the temporal structure, such as a reduction of the reaction time (RT; Correa & Nobre, 2008; Sanabria, Capizzi, & Correa, 2011; Sanabria & Correa, 2013), time judgement (Barnes and Jones, 2000) or pitch judgment accuracy (Jones, Moynihan, MacKenzie, & Puente, 2002).

This synchronisation with external rhythms, known as rhythmic entrainment, is not only restricted to auditory stimuli. Several studies have found entrainment, resulting in behavioural benefits, using both dynamic visual stimuli (e.g., a ball moving across the screen; Doherty, Rao, Mesulam & Nobre, 2005; Rohenkol, Coull & Nobre, 2011) and sequences of stationary visual stimuli (Marchant and Driver, 2013). However, previous research has revealed behavioural advantages of auditory over visual rhythms, showing that stimulus duration is more accurately perceived in the auditory modality, as measured by discrimination and judgement tasks (Glenberg & Jona, 1991; Repp & Penel, 2002; Guttman, Gilroy, & Blake, 2005). Auditory dominance for temporal information processing has been shown, for example, in studies where a repetitive sound is presented simultaneously with a flickering light (e.g., Welch, Dutton Hurt, & Warren, 1986). In line with these studies, one could expect a stronger entrainment for the auditory rhythms. However, to our knowledge, no study has directly compared auditory with visual rhythmic entrainment.

Regardless of the sensory modality in which the stimuli are presented, several studies have focused on the role of attention in rhythmic entrainment. It has been shown, for example, that participants' responses to target stimuli presented after an isochronous rhythm are faster than those presented after an anisochronous rhythm (e.g., stimuli presented at random intervals), even when participants are specifically instructed to ignore the rhythm (Rohenkohl, Coull, & Nobre, 2011). Results like this suggest that attentional focus is not necessary for the rhythm to produce entrainment.

In this line, De la Rosa and colleagues (De la Rosa, Sanabria, Capizzi, & Correa, 2012) tested the resistance of auditory rhythmic entrainment under a dual-task paradigm, engaging the rhythmic entrainment process in competition for attentional resources using a secondary task. In these

experiments, participants had to perform an RT task concurrently with a working memory (WM) Sternberg-type task, in which participants had to remember a set of letters presented at the beginning of each trial. After the RT task, a single letter was presented and participants had to respond by indicating whether this letter was presented or not in the original set. Results showed that participants' RTs were faster after the isochronous sequences than after the anisochronous sequences, and that the secondary task had no influence on this effect.

More recently, Cutanda and colleagues (Cutanda, Correa & Sanabria, 2015) went a step further to test the resistance of the entrainment processes under more demanding conditions of the secondary task. Unlike in the De la Rosa et al. (2012) study, all tasks were presented in the same modality (auditory) to increase resource competition. Moreover, the secondary task, performed concurrently with the RT task, required an updating in WM. We specifically used an n-back task in which a single letter was presented on each trial and participants had to respond by indicating whether this letter matched the one presented *n* trials earlier. Results showed an overall increase of the RT for the 2-back condition, showing that the dual-task manipulation was effective, but, once again, the entrainment effect (faster RT after isochronous than after anisochronous sequences) was not affected by the task, since the RT benefit induced by the isochronous rhythm was similar for both the high-load (2-back) and the low-load (0-back) conditions. If there were competition for limited resources common to both tasks, one would have expected that the addition of the secondary task would interfere with the primary task (Logan, 1979). The resistance of the entrainment effect to this interference therefore suggested that rhythm entrainment is an automatic process.

The manipulation of regularity in previous studies has usually included two extreme conditions, perfect isochrony and full anisochrony (e.g. Lange, 2010; Rohenkohl, Coull, & Nobre, 2011; De la Rosa et al., 2012; Cutanda et al., 2015). Results arising from these designs can only provide dichotomic information regarding the automatic nature of rhythmic entrainment, depending on whether this effect is resistant or not to an attention- demanding secondary task. However, automaticity of cognitive processing can be understood as a continuum (MacLeod & Dunbar, 1988; Capizzi, Sanabria, & Correa, 2012).

With the purpose of overcoming the limitation of previous designs, in the present study we used a parametric manipulation of the isochrony in the stimulus sequences presented before the target. The sequences could be either isochronous or anisochronous, the latter ranging from either smaller to larger deviations of the isochrony and, thus, resulting in a range of different anisochronous sequences. This manipulation had two objectives: 1) to obtain

and compare the entrainment patterns for both auditory and visual modalities (Experiment 1); and 2) to test the resistance of rhythmic entrainment to a WM secondary task (Experiment 2).

In Experiment 1, participants were randomly assigned to either an auditory or visual rhythm condition. We expected to observe a progressive slowing of the RT as the deviations from isochrony increased (i.e., "the entrainment effect"), that is, we analysed participants' RTs to detect a target occurring at a fixed interval as a function of the level of the anisochrony of the entraining rhythm. We further expected this entrainment effect to be larger for auditory than for visual rhythms.

EXPERIMENT 1

METHOD

Participants. Participants gave written consent before taking part in the experiment in return for monetary compensation (10 Euros). In Experiment 1, twenty-four students (11 men) from the University of Granada, aged between 20 and 24 years (mean age = 20.75, SD = 1.22), were pseudo-randomly (alternatively) assigned to the auditory group. Five of the participants reported having formal musical training for at least the last 24 months prior to the experiment. Twenty-four students (10 men) from the University of Granada, with ages between 19 and 25 years (mean age = 20.88; SD= 1.62), were assigned to the visual group. None of the participants reported having formal musical training in the 24 months prior to the experiment.

None of the participants reported any history of neuropsychological disorders or uncorrected hearing and visual impairments. The experiment was conducted according to the ethical standards set out in the Declaration of Helsinki (1964) and approved by the ethics committee of the University of Granada (Ref.: 17/CEIH/2015).

Apparatus and stimuli. The experiment was run on an Intel Core i7 computer connected to a 19-in. LCD monitor with a refresh rate of 100Hz. The E-prime software (Schneider, Eschman, & Zuccolotto, 2002) was used for presentation of the stimuli and to record participants' responses. The viewing distance was approximately 60 cm. In the auditory task, participants had to respond to a target tone with a frequency of 400 Hz and 50 ms duration. This stimulus was preceded by an auditory sequence of 17 tones of 700 Hz and 50 ms duration. The frequencies of each tone in the rhythmic sequences

as well as the frequency of the target tone were the same as in previous studies (Cutanda, Correa, & Sanabria, 2015; De la Rosa, Sanabria, Capizzi, & Correa, 2012; Sanabria, Capizzi, & Correa, 2011). All auditory stimuli were presented via headphones at a clearly audible sound intensity. In the visual task, the stimuli in the rhythmic sequence prior to the target consisted of a set of 17 identical visual stimuli: the letter "O" presented in the centre of the screen for 50 ms, in uppercase, 280-point Courier New, black colour font. The target to which participants had to respond, presented after this sequence, was a 50 ms black and white 96 x 96 pixels Gabor patch. All stimuli were presented on a grey background.

Procedure and task. Participants were tested in a silent and dimly illuminated room. They were provided with both written and spoken instructions to become familiar with the stimuli and task. They were informed that a sequence of tones (auditory task) or visual stimuli (visual task) irrelevant to the task would be presented before the target. To make sure that participants could clearly recognise the target, they could also listen to it or see it as many times as they wanted before performing the task. Each trial began with the presentation of a fixation point (plus sign, 1.5° x 1.5°) in the centre of the screen, which remained on and steady until the target presentation at the end of each trial. After 500 ms, the sequence was presented, and 17 stimuli (each of 50 ms in duration) were presented on each trial. For each of the seven conditions, a value of 0, 10, 20, 50, 100, 150, or 200 ms was either added or subtracted to 300 ms at each interval within the sequence. In this way, rhythms were produced which could range between isochrony (0-ms deviation) to the largest anisochrony (200-ms deviation) (See Figure 1). For each condition, one of these values was randomly added to 300 ms in half of the intervals and subtracted in the remaining half so that all the sequences had the same duration. Since the inclusion of more than one foreperiod (i.e. one short and one long foreperiod) would result in a foreperiod effect (faster responses after the longest than after shorter foreperiods; Sanabria et al., 2011), only a fixed foreperiod of 650 ms was presented after the sequence and prior to the target onset. The foreperiod duration was consistent with the temporal structure of the fully isochronous sequence, since it was twice as long as the interval between each stimulus, plus the duration of one of the stimuli that would occur between intervals.

All trials had a total duration of 6350 ms from the first stimulus of the sequence to the target onset. Participants responded as quickly as possible to the target by pressing the spacebar key with their right index finger. A feedback message informing participants of their RTs was displayed for 500 ms and also if no response was given within 1000 ms after the target onset ("respond faster").

Each participant completed 1 practice and 6 experimental blocks of 35 trials each, resulting in 30 trials per experimental condition in both auditory and visual tasks. The experiment had a total duration of approximately 45 minutes.
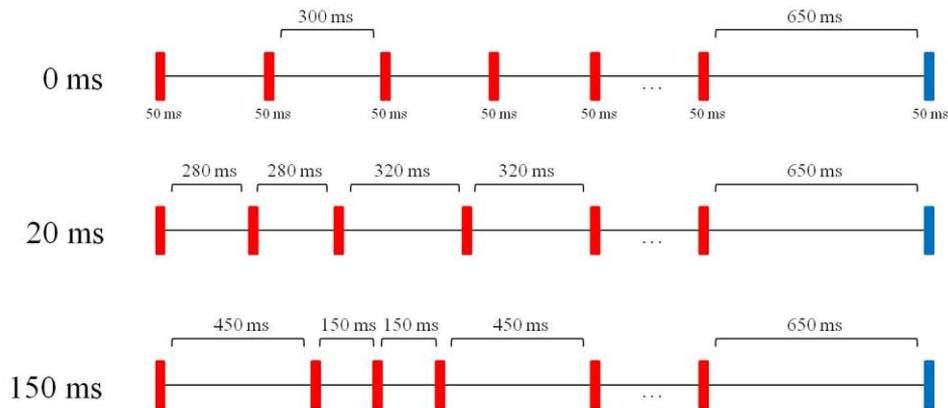

Fig 1. Schematic representation of events in Experiment 1, with examples of three of the seven anisochrony conditions.

Design and data analysis. The mean RT to respond to the target was obtained for each one of the seven anisochrony conditions (0, 10, 20, 50, 100, 150 and 200 ms). In order to compare rhythmic entrainment between visual and auditory modalities, we performed between-groups analyses for the RT, including modality (visual and auditory) as between-participants factor and anisochrony (0, 10, 20, 50, 100, 150 and 200 ms) as the within-participants factor. Data were analysed by a two-way analysis of variance (ANOVA).

Data from the practice block and trials in which participants responded before the target onset (during the foreperiod), or in which no response was given, were not included in the analysis (auditory group, 7.1%; visual group, 5.75%). Premature responses (responses given during the first 100 ms after the target onset) were also excluded from the RT analysis to ensure the inclusion of participants' responses in reaction to the target (auditory group, 19.3%; visual group, 12.5%). Due to the unexpected high amount of premature responses, we conducted an exploratory analysis on their proportions following the same design (modality x anisochrony).

Significant main effects and interactions were further analysed by multiple paired t-test and corrected by Holm-Bonferroni test for multiple comparisons. Greenhouse-Geisser's corrected p-values and degrees of freedom are reported when the Mauchly's test indicated a violation of sphericity. Cohen's d was obtained for each comparison as well as partial eta-

squared for the size of main effects and interactions. The 0.95 confidence intervals were also computed and reported in the figures.

RESULTS

For the ANOVA conducted on the participants' mean RT, sphericity was violated, as indicated by Mauchly's test, $(X^2(20) = 52.24, p < .001)$; therefore, the degrees of freedom were corrected by Greenhouse-Geisser ($\varepsilon = 0.72$). As expected, results showed a significant main effect of modality, $F(1, 46) = 6.7$, $p=.013$, $\eta_\rho^2=.13$, revealing that participants responded faster in the auditory condition (M = 228 ms) than in the visual condition (M = 261 ms). The main effect of anisochrony was also significant, $F(4.29, 197.51)= 60.68$, $p<.001$, $\eta_\rho^2=.57$. Holm-Bonferroni test revealed no differences between the 0, 10, 20 and 50 ms conditions (i.e. the "entrainment window", see Figure 2). For larger anisochronies, the RT increments of the 100, 150 and 200 ms conditions with respect to the 0 ms condition reached significance (all $ps<.001$). RTs further increased progressively after the 100 ms anisochrony, showing significant differences between 100 and 150 ms, and 150 and 200 ms (both $ps<.001$). (See Table 1, Supplementary material for further details). Interestingly, the interaction between modality and anisochrony was far from significant, $F < 1$.
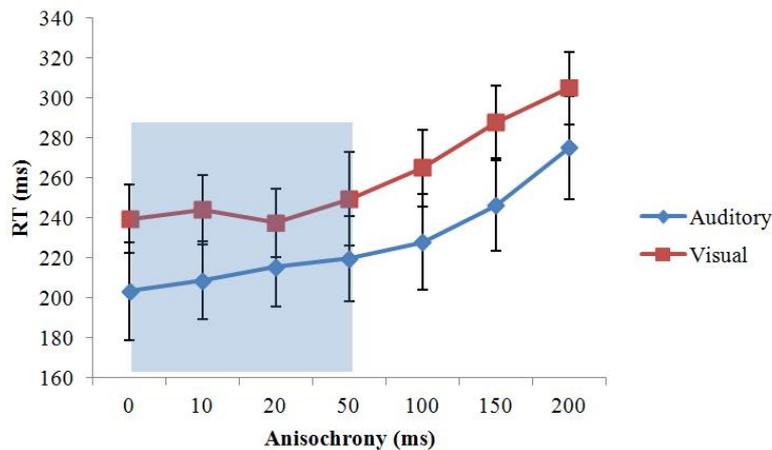

Fig 2. Mean reaction times (RT) as a function of anisochrony (0, 10, 20, 50, 100, 150 and 200 ms) and modality (auditory and visual) in Experiment 1. Shaded area indicates the entrainment window where no significant differences were shown between anisochrony conditions. Vertical bars represent 95% confidence intervals.

The ANOVA on the proportion of premature responses showed a significant main effect of modality, $F(1, 46) = 6.29$, $p=.016$, $\eta_\rho^2=.12$, with more premature responses for the auditory (M = .19) than for the visual condition (M = .12). The main effect of anisochrony was also significant, $F(6, 276) = 41.85$, $p<.001$, $\eta_\rho^2=.48$ (see Figure S1, Supplementary material). Holm-Bonferroni test revealed a larger proportion of premature responses in the 0 ms condition in comparison with the 100, 150 and 200 ms conditions (all $ps<.001$), with no significant differences between the 0 ms and the 10, 20 and 50 ms conditions. Following a similar pattern to that of the RT, the number of premature responses progressively decreased after the 100 ms anisochrony, showing significant differences between the 100 ms and the 150 ms conditions and the 150 ms and the 200 ms conditions respectively. The interaction between modality and anisochrony was significant, $F(6, 276) = 6.63$, $p<.001$, $\eta_\rho^2=.13$, but the Holm-Bonferroni test revealed significant differences between modalities only for the 10 ms condition ($p=.033$).

DISCUSSION

The analysis of Experiment 1 showed an entrainment window (a peak of the entrainment effect, with faster RTs) between the 0-ms and anisochronies up to 50 ms. After this entrainment window, RTs became progressively slower as the anisochronies increased. These results do not only point to the gradual nature of rhythmic entrainment, but also suggest that sequences with a relatively high degree of anisochrony can also lead to entrainment.

The RT analysis did not show a significant interaction between modality and anisochrony, but the analysis with premature responses did reveal such an interaction. However, the Holm-Bonferroni test revealed that these differences were only significant for the 10 ms condition, so we cannot conclude that the auditory modality leads to a stronger entrainment effect. Rather, it seems that both modalities lead to a similar pattern of entrainment, triggering faster responses after the more isochronous sequences regardless of the modality in which the stimuli were presented.

Experiment 1 allowed us to observe a gradual pattern of rhythmic entrainment, and to directly compare auditory and visual entrainment effects. In Experiment 2, we added a secondary working memory task to the procedure presented in Experiment 1, with the purpose of testing the resistance of both auditory and visual rhythmic entrainment under dual-task conditions.

EXPERIMENT 2

METHOD

Participants. As in Experiment 1, here participants also gave written consent before taking part in the experiment in return for monetary compensation (10 Euros). Twenty-four students (8 men) from the University of Granada, with ages between 19 and 29 years (mean age = 20.92; SD= 2.21), were pseudo-randomly assigned to the auditory group. Twenty-four students (11 men) from the University of Granada, with ages between 19 and 34 years (mean age = 20.96; SD= 3.17), were assigned to the visual group. Four of the participants in the auditory group and two in the visual group reported having formal musical training in the 24 months prior to the experiment. Their results did not differ from the rest of the participants and their inclusion or exclusion from the analysis did not significantly change the results.

None of the participants reported any history of neuropsychological disorders or uncorrected hearing and visual impairments. The experiment was conducted according to the ethical standards set out in the Declaration of Helsinki (1964) and was approved by the ethics committee of the University of Granada.

Apparatus and stimuli. The hardware, software and stimuli for the rhythmic entrainment task were the same as in Experiment 1. For the WM task, four letters (M, R, S and V) were chosen for the N-Back task. One of the letters was randomly presented at the beginning of each trial. In the auditory task all letters were orally presented by a male voice, which had been previously recorded and edited using Audacity 2.0.2 software (The Audacity Team, 2012). All auditory stimuli were presented at the same intensity as in Experiment 1. In the visual task, the letters were visually presented in uppercase, 18-point Courier New, black colour font. In both experiments all letters were presented for 500 ms.

Procedure and task. For the rhythmic entrainment task the procedure was the same as in Experiment 1. For the N-back task, participants were provided with both spoken and written instructions were also familiarised with the stimuli before performing the task. Each trial began with the presentation of a fixation point (plus sign, 1.5° X 1.5°) for 500 ms, followed by either the oral (auditory task) or visual (visual task) presentation of a letter for 500 ms, depending on the modality group condition. After this, a question mark appeared in the centre of the screen for 3000 ms (see Figure 3). At this point, participants were asked to press the "a" key if the letter was the same

as the one presented on the first trial of the block (0-back blocks) or two trials ago (2-back blocks), and the "z" key if the letter was different. Feedback on the WM performance was presented after the response for 300 ms. The rhythmic entrainment task was then presented, this being the same as in Experiment 1. Stimulus modality (visual or auditory) was always the same for both WM and rhythmic entrainment tasks. The presentation order of the 0-back and the 2-back task blocks was counterbalanced across participants, with half of them responding first to the 0-back task and the other half to the 2-back task. Each trial had a total duration of 10150 ms.

Each participant completed 1 practice block and 3 experimental blocks of 35 trials each in both WM load conditions, resulting in 15 trials per experimental condition in both auditory and visual tasks. The experiment had a total duration of approximately 75 minutes.
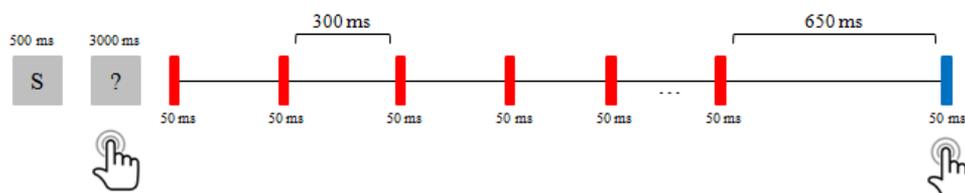

Fig 3. Schematic representation of events in one trial of Experiment 2 for the 0 ms anisochrony condition.

Design and data analysis. The experiment constituted a 2x7x2 design with the independent variables of modality (auditory and visual) as between-participants factor and anisochrony (0, 10, 20, 50, 100, 150 and 200 ms deviations) and memory load (0-back and 2-back) as within-participants factors. From here we conducted two separate ANOVAS with the dependent variables of mean RT to the target and accuracy in the WM task. Finally, in light of the results found in Experiment 1 we conducted an ANOVA with the proportion of premature responses as a dependent variable.

As in Experiment 1, data from the practice block and trials in which participants responded during the foreperiod or without response (auditory group, 6.05%; visual group, 4.52%) and premature responses (auditory group, 9.38%; visual group, 11.23%) were not included in the RT analysis. To ensure that participants were engaged in the WM task, trials with incorrect responses in the WM task were also discarded for the analysis of the mean RT.

As in Experiment 1, multiple paired t-tests corrected by Holm-Bonferroni test for multiple comparisons were run to further analyse

significant main effects and interactions. Greenhouse-Geisser's corrected p-values and degrees of freedom are reported when Mauchly's test indicated a violation of sphericity. Cohen's d for each comparison as well as partial eta-squared for the size of main effects and interactions are reported.

RESULTS

For The analysis of the accuracy data in the WM task was first conducted to check whether the dual-task manipulation was effective. Indeed, the results showed a significant main effect of memory load, $F(1, 46)= 79.51$, $p<.001$, $\eta_\rho^2=.63$, with higher accuracy on the 0-back (M = .95) than the 2-back (M = .85) task. The main effect of modality was also significant, $F(1, 46)= 10.33$, $p=.002$, $\eta_\rho^2=.18$, with participants being more accurate in the visual (M = .93) than the auditory (M = .88) modality. Finally, there was a significant interaction between memory load and modality, $F(1, 46)= 5.95$, $p=.019$, $\eta_\rho^2=.11$, showing a larger difference in accuracy between modality conditions in the 2-back (mean auditory: .82; mean visual: .89) than in the 0-back task (mean auditory: .94; mean visual: .96). None of the other terms reached statistical significance (anisochrony, $F < 1$; anisochrony x modality, $F(6, 276)= 1.2$, $p=.31$, $\eta_\rho^2=.03$; memory load x anisochrony, $F(6, 276)= 1.19$, $p=.31$, $\eta_\rho^2=.03$; memory load x anisochrony x modality, $F < 1$; see Figure 3).
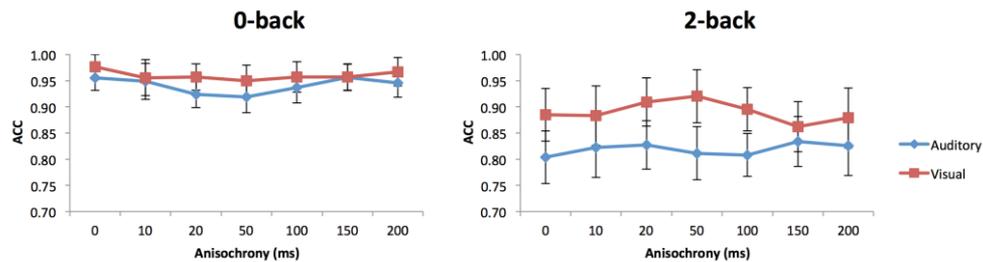
Fig 3. Mean proportion of accuracy on the n-back task as a function of anisochrony (0, 10, 20, 50, 100, 150 and 200 ms), modality (auditory and visual) and memory load (0-back and 2-back) in Experiment 2. Vertical bars represent 95% confidence intervals.

Sphericity was not violated in the ANOVA with participants' RTs, as indicated by Mauchly's test, $(X^2(20) = 17.86, p = .59)$. The analysis showed a significant main effect of anisochrony, $F(6, 276)= 31.4$, $p<.001$, $\eta_\rho^2=.41$. As in Experiment 1, Holm-Bonferroni test showed no differences between the 0 ms condition and the 10, 20 and 50 ms conditions. The RT increments for larger anisochronies with respect to the isochronous (0 ms) condition were

significant (100 ms, $p = .002$; 150 ms and 200 ms, both $p < .001$). Finally, the results revealed significant differences between the 200 ms and the 100 ms and 150 ms conditions (both $ps < .001$). There were no significant differences between the 0 ms condition and the 10 ms, 20 ms and 50 ms conditions (see Table 2, Supplementary material for further details). The main effect of memory load was also significant, $F(1, 46)= 16.17$, $p<.001$, $\eta_p^2=.26$, with faster RTs in the 0-back task (M = 285 ms) than in the 2-back task (M = 307 ms), but this did not interact with anisochrony, $F(6, 276)= .43$, $p=.86$, $\eta_p^2=.009$. None of the other terms in the ANOVA reached statistical significance (all $Fs < 1$), in line with Figure 4, which shows a similar pattern of entrainment for both the 0-back and the 2-back tasks, and for both auditory and visual modalities.

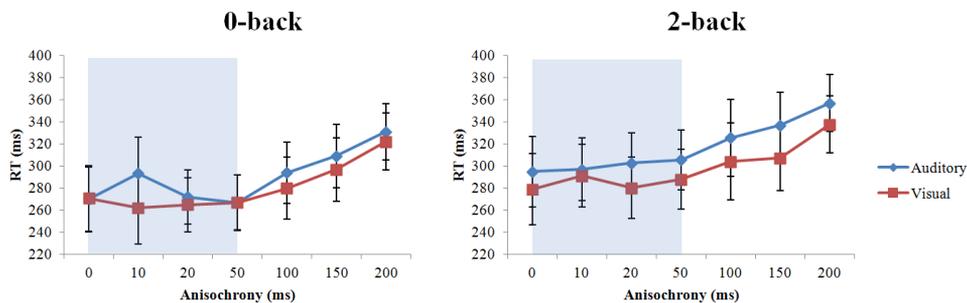**Fig 4**. Mean reaction times (RT) as a function of the anisochrony (0, 10, 20, 50, 100, 150 and 200 ms), modality (auditory and visual) and memory load (0-back and 2-back) in Experiment 2. Shaded area indicates the entrainment window where no significant differences were found between anisochrony conditions. Vertical bars represent 95% confidence intervals.

The ANOVA on the premature responses showed a main effect of anisochrony, $F(6, 276)= 18.24$, $p<.001$, $\eta_p^2=.28$, with significant differences between the 0 ms and the 100, 150 and 200 ms conditions (all $ps<.001$), and no differences between the 0 ms and the 10, 20 and 50 ms anisochronies. Significant differences were also found between the 100 ms and the 200 ms condition ($p=.02$) but not between 150 and 200 ms (see Figure S2, Supplementary material). Main effects of memory load and modality were not significant ($F=2.63$ and $F<1$, respectively). None of the interactions were significant (Anisochrony x Memory Load, $F=1.33$; Anisochrony x Modality, Modality x Memory Load, Anisochrony x Modality x Memory Load, all $Fs < 1$).

DISCUSSION

Research has shown that auditory rhythmic entrainment results in behavioural benefits such as faster RT (Sanabria, Capizzi, & Correa, 2011; Sanabria & Correa, 2013) or better time and pitch discrimination (Barnes and Jones, 2000; Jones, Moynihan, MacKenzie, & Puente, 2002), and that the entrainment processes can also be observed in the visual modality (Doherty, Rao, Mesulam & Nobre, 2005; Rohenkol, Coull & Nobre, 2012; Marchant and Driver, 2013). Moreover, rhythmic entrainment has proven to be resistant to demanding secondary tasks (De la Rosa, Sanabria, Capizzi, & Correa, 2012; Cutanda, Correa, & Sanabria, 2015). Taken together, these results suggest that rhythmic entrainment is a strong, highly automatic process, and it can be triggered by both auditory and visual rhythmic stimuli.

However, two main issues can be drawn from the mentioned studies. First, entrainment had only been studied under extreme dichotomic conditions, thus precluding a systematic analysis of its potentially gradual nature. Second, to our knowledge, visual and auditory entrainment had not been compared directly and, thus, it was unknown as to whether there is an advantage of one modality over the other. Two experiments were designed to address these issues, studying the potential differences between auditory and visual modalities in rhythmic entrainment (Experiment 1), and testing the resistance of rhythmic entrainment under dual-task conditions (Experiment 2). The main finding of the current research shows that this entrainment effect (i.e. faster responses for more isochronous sequences) was similarly present for both auditory and visual rhythmic cues, and for both low- and high-load working memory tasks.

The analysis of both Experiments 1 and 2 confirmed that participants were faster at the smaller than at the larger deviations from isochrony. The RT results of the present study showed two different periods of entrainment as a function of anisochrony. First, the peak of the entrainment effect (fastest RTs) was observed within a 50-ms window including the three smallest deviations (10, 20 and 50 ms conditions) in addition to the full isochrony condition. Second, beyond this "entrainment window", participants' RTs became progressively slower as the deviations from isochrony increased, suggesting a progressive disentrainment. These results suggest that sequences with small anisochronies can also lead to entrainment effects as strong as pure isochronic sequences.

In line with several studies showing an auditory advantage in temporal processing (Glenberg & Jona, 1991; Repp & Penel, 2002; Guttman, Gilroy, & Blake, 2005; Grahn, 2013), we expected differences in the RT entrainment function between auditory and visual stimuli. Moreover, visuo-motor

synchronisation has proven to be more stable than audio-motor synchronisation when motion stimuli are used, whereas stationary stimuli seem to benefit audio-motor synchronisation (Hove, Fairhurst, Kotz, & Keller, 2013). Given that all the stimuli in Experiments 1 and 2 were stationary, one could expect to find a stronger entrainment effect with auditory rhythms.

In fact, the finding of an increased proportion of premature responses in the auditory (vs. visual) condition was congruent with a more precise motor synchronisation to the stationary auditory stimuli than to the visual sequences. Premature responses also revealed an entrainment effect as a function of isochrony, suggesting that rhythms could have helped to estimate and anticipate the time of target onset (Barnes & Jones, 2000). This entrainment effect on temporal anticipation is consistent with research linking temporal preparation with harder response inhibition of prepotent responses (Correa et al., 2010; Korolczuk, Burle, & Coull, 2018); that is, temporal anticipation decreased as the anisochrony increased, so that participants progressively needed less inhibitory control to wait for and respond to the actual target onset.

It should be noted, however, that the exploratory analysis on the premature responses was conducted due to the unexpected high proportion of participants' RT under 100 ms. Although these results could provide further information about the potential differences between auditory and visual entrainment, our main results for the RT suggest that the entrainment process is not affected by the modality in which the sequences are presented, since the isochrony of the sequences led to faster responses and followed the same pattern regardless of the modality in which these sequences were presented. Moreover, the interaction between modality and anisochrony for the premature responses invites further replication, since Holm-Bonferroni analysis only showed differences in the 10 ms condition. More experiments comparing both auditory and visual entrainment should be conducted with dynamic stimuli to test whether there could be a higher proportion of premature responses as well as faster RTs for the visual than for the auditory stimuli.

In Experiment 2, no significant differences were found in rhythmic entrainment as a function of the memory load, in line with our previous study (Cutanda, Correa, & Sanabria, 2015). Moreover, the entrainment window in Experiment 2 was the same as that in Experiment 1. Taken together, these results suggest a strong resistance of the entrainment effect to the simultaneous performance of a secondary task, pointing once again to a high degree of automaticity. These results are in line with those of previous studies showing that top-down attention does not play an important role in rhythmic

entrainment (Bégel et al., 2017; Correa et al., 2014; Rohenkohl, Coull, & Nobre, 2011). In this same context, recent research on auditory predictive processing suggests that attentional control is not required to generate predictions (see Bendixen, SanMiguel, & Schröger, 2012, for a review). This would explain the lack of an influence of the secondary task on the entrainment effect observed in Experiment 2, which also supports the idea that visual isochronous sequences can generate temporal predictions as strongly as auditory sequences, since the dual-task had no influence on either our auditory or visual tasks.

Several authors have considered that neural oscillations could be the mechanism by which regular rhythms improve behavioural performance (Lakatos et al., 2008; Schroeder & Lakatos, 2009; for reviews, see Arnal & Giraud, 2012, and Nobre & van Ede, 2018). Results suggest an entrainment of neural oscillations to an external isochronous rhythm, synchronising (phase-locking) with the pattern of the temporal structure and, in consequence, aligning the high-excitability phases of the brain dynamics with the expected stimuli. According to these studies, this synchronisation would occur when the rhythm is detected in a task, and would thus explain why our participants showed an entrainment effect not only in the auditory tasks, but also in the visual tasks. Interestingly, it has been shown that isochronous visual sequences can also entrain neural oscillation, phase-locking to the visual stimuli (Mathewson et al., 2012). Finally, and regardless of neural oscillations, it has also been shown that temporal predictability of a visual isochronous sequence results in enhanced activity of several structures previously reported in auditory temporal processing (dorsolateral prefrontal cortex, left intraparietal sulcus, inferior frontal gyrus, insula and basal ganglia; Marchant & Driver, 2013). These findings, together with our RT results, suggest that the entrainment of neural oscillations to these rhythmic sequences could result in a temporal prediction of when the target is going to appear, and that this prediction could occur in both auditory and visual domains. The participants' mean RT in both experiments were used to measure the effect of the sequences on entrainment; however, entrainment can also result in more stable responding to the target, with a reduction in the trial-by-trial variability of the RT (Steinborn, Langner, & Huestegge, 2017). Further analysis regarding variability could provide additional support for the results found in Experiments 1 and 2.

In summary, the results of both Experiments 1 and 2 confirmed that rhythmic entrainment could be triggered not only by purely isochronous stimuli, but also by rhythmic sequences with a certain degree of anisochrony. Moreover, rhythmic entrainment has proven to be resistant to demanding secondary tasks and did not show any gradation of automaticity. Our research

suggests that rhythmic entrainment constitutes a strong, and probably basic, cognitive process that is not dependent on either top-down attention or on the modality in which the stimuli are presented. This process would instead occur by default, generating behavioural benefits in response to events matching a temporal structure.

RESUMEN

La anisocronía de una secuencia de estímulos se ha manipulado paramétricamente para investigar si la sincronización rítmica es más potente en la modalidad auditiva que en la visual (Experimento 1), y si esta depende de procesos de la atención de tipo arriba-abajo (Experimento 2). En el Experimento 1, los participantes tenían que responder lo más rápido posible ante un estímulo objetivo presentado después de una secuencia de estímulos que podían ser visuales o auditivos. La anisocronía de esta secuencia era manipulada paramétricamente en lugar de siguiendo un procedimiento discreto del tipo "todo o nada"; es decir, la anisocronía podía variar en un rango de pequeñas a mayores desviaciones (0, 10, 20, 50, 100, 150 y 200 ms). Los patrones resultantes de la sincronización rítmica fueron comparados para las modalidades auditiva y visual. Los resultados mostraron un pico de sincronización para las condiciones de isocronía y las condiciones de anisocronía que tenían hasta un máximo de 50 ms de desviación (es decir, los participantes respondían igual de rápido para las condiciones de desviación 0, 10, 20 y 50 ms), lo cual sugiere que las secuencias anisócronas también pueden producir sincronización rítmica. A partir de esta ventana de sincronización, los tiempos de reacción fueron progresivamente más altos. Sorprendentemente, no encontramos diferencias en los patrones de sincronización entre los ritmos auditivos y los visuales.En el Experimento 2, utilizamos una metodología de tarea dual mediante la inclusión de una tarea n-back de memoria de trabajo al procedimiento del Experimento 1. Los resultados no mostraron una interferencia de la tarea secundaria ni en la modalidad auditiva ni en la visual, pues los participantes mostraron el mismo patrón de sincronización que en el Experimento 1. Estos resultados sugieren que la sincronización rítmica constituye un proceso cognitivo que beneficia el comportamiento, que ocurre por defecto (automáticamente),

independientemente de la modalidad por la que se presenten los estímulos, y que es independiente de procesos "arriba-abajo" de la atención.

REFERENCES

Arnal, L. H., & Giraud, A.L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390-398. https://doi.org/10.1016/j.tics.2012.05.003

Barnes, R., & Jones, M. R. (2000). Expectancy, Attention, and Time. *Cognitive Psychology*, *41*(3), 254-311. https://doi.org/10.1006/cogp.2000.0738

Bégel, V., Benoit, C.E., Correa, A., Cutanda, D., Kotz, S. A., & Dalla Bella, S. (2017). "Lost in time" but still moving to the beat. *Neuropsychologia*, *94*, 129–138. https://doi.org/10.1016/j.neuropsychologia.2016.11.022

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: a review. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *83*(2), 120-131. https://doi.org/10.1016/j.ijpsycho.2011.08.003

Brown, S. W. (2006). Timing and executive function: Bidirectional interference between concurrent temporal production and randomization tasks. *Memory and Cognition, 34 (7),* 1464–1471. https://doi.org/10.3758/BF03195911

Capizzi, M., Sanabria, D., & Correa, Á. (2012). Dissociating controlled from automatic processing in temporal preparation. *Cognition*, *123*(2), 293-302. https://doi.org/10.1016/j.cognition.2012.02.005

Correa, Á., Cona, G., Arbula, S., Vallesi, A., & Bisiacchi, P. (2014). Neural dissociation of automatic and controlled temporal preparation by transcranial magnetic stimulation. *Neuropsychologia*, *65*, 131–136. https://doi.org/10.1016/j.neuropsychologia.2014.10.023

Correa, Á., & Nobre, A. C. (2008). Neural Modulation by Regularity and Passage of Time. *Journal of Neurophysiology*, *100*(3), 1649-1655. https://doi.org/10.1152/jn.90656.2008

Correa, Á., Triviño, M., Pérez-Dueñas, C., Acosta, A. & Lupiáñez, J. (2010). Temporal preparation, response inhibition and impulsivity. Brain and Cognition, 73, 222-228. https://doi.org/10.1016/j.bandc.2010.05.006

Cutanda, D., Correa, Á., & Sanabria, D. (2015). Auditory temporal preparation induced by rhythmic cues during concurrent auditory working memory tasks. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(3), 790-797. https://doi.org/10.1037/a0039167

de la Rosa, M. D., Sanabria, D., Capizzi, M., & Correa, Á. (2012). Temporal Preparation Driven by Rhythms is Resistant to Working Memory Interference. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00308

Doherty, J. R., Rao, A., Mesulam, M. M., & Nobre, A. C. (2005). Synergistic Effect of Combined Temporal and Spatial Expectations on Visual Attention. *The Journal of Neuroscience*, *25*(36), 8259-8266. https://doi.org/10.1523/JNEUROSCI.1821-05.2005

Fortin, C., & Breton, R. (1995). Temporal interval production and processing in working memory. *Perception & Psychophysics, 57*, 203–215. https://doi.org/10.3758/BF03206507

Gan, L., Huang, Y., Zhou, L., Qian, C., & Wu, X. (2015). Synchronization to a bouncing ball with a realistic motion trajectory. *Scientific Reports*, *5*, 11974. https://doi.org/10.1038/srep11974

Glenberg, A. M., & Jona, M. (1991). Temporal coding in rhythm tasks revealed by modality effects. *Memory & Cognition*, *19*(5), 514-522. https://doi.org/10.3758/BF03199576

Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research*, *220*(1), 51-61. https://doi.org/10.1007/s00221-012-3114-8

Guttman, S. E., Gilroy, L. A., & Blake, R. (2005). Hearing What the Eyes See. *Psychological science*, *16*(3), 228-235. https://doi.org/10.1111/j.0956-7976.2005.00808.x

Hove, M. J., Fairhurst, M. T., Kotz, S. A., & Keller, P. E. (2013). Synchronizing with auditory and visual rhythms: An fMRI assessment of modality differences and modality appropriateness. *NeuroImage*, *67*, 313-321. https://doi.org/10.1016/j.neuroimage.2012.11.032

Hove, M. J., Spivey, M. J., & Krumhansl, C. L. (2010). Compatibility of motion facilitates visuomotor synchronization. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(6), 1525-1534. https://doi.org/10.1037/a0019059

Iversen, J. R., Patel, A. D., Nicodemus, B., & Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition*, *134*, 232-244. https://doi.org/10.1016/j.cognition.2014.10.018

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal Aspects of Stimulus-Driven Attending in Dynamic Arrays. *Psychological Science*, *13*(4), 313-319. https://doi.org/10.1111/1467-9280.00458

Korolczuk, I., Burle, B., & Coull, J. T. (2018). The costs and benefits of temporal predictability: impaired inhibition of prepotent responses accompanies increased activation of task-relevant responses. *Cognition*, *179*, 102-110. https://doi.org/10.1016/j.cognition.2018.06.006

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science*, *320*(5872), 110-113. https://doi.org/10.1126/science.1154735

Lange, K. (2010). Can a regular context induce temporal orienting to a target sound? *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *78*(3), 231-238. https://doi.org/10.1016/j.ijpsycho.2010.08.003

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 189–207. http://dx.doi.org/10.1037/0096-1523.5.2.189

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology*, *14*, 126-135. http://dx.doi.org/10.1037/0278-7393.14.1.126

Marchant, J. L., & Driver, J. (2013). Visual and Audiovisual Effects of Isochronous Timing on Visual Perception and Brain Activity. *Cerebral Cortex (New York, NY)*, *23*(6), 1290-1298. https://doi.org/10.1093/cercor/bhs095

Mathewson, K. E., Prudhomme, C., Fabiani, M., Beck, D. M., Lleras, A., & Gratton, G. (2012). Making Waves in the Stream of Consciousness: Entraining Oscillations in EEG Alpha and Fluctuations in Visual Awareness with Rhythmic Visual Stimulation. *Journal of Cognitive Neuroscience*, *24*(12), 2321-2333. https://doi.org/10.1162/jocn_a_00288

Nobre, A., Correa, A., & Coull, J. (2007). The hazards of time. *Current Opinion in Neurobiology*, *17*(4), 465-470. https://doi.org/10.1016/j.conb.2007.07.006

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nature Reviews Neuroscience*, *19*(1), 34–48. https://doi.org/10.1038/nrn.2017.141

Provasi, J., & Bobin-Bègue, A. (2003). Spontaneous motor tempo and rhythmical synchronisation in 2½- and 4-year-old children. *International Journal of Behavioral Development*, *27*(3), 220-231. https://doi.org/10.1080/01650250244000290

Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology. Human Perception and Performance*, *28*(5), 1085-1099. http://dx.doi.org/10.1037/0096-1523.28.5.1085

Rohenkohl, G., Coull, J. T., & Nobre, A. C. (2011). Behavioural Dissociation between Exogenous and Endogenous Temporal Orienting of Attention. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0014620

Sanabria, D., Capizzi, M., & Correa, Á. (2011). Rhythms that speed you up. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(1), 236-244. https://doi.org/10.1037/a0019956

Sanabria, D., & Correa, Á. (2013). Electrophysiological evidence of temporal preparation driven by rhythms in audition. *Biological Psychology*, *92*(2), 98-105. https://doi.org/10.1016/j.biopsycho.2012.11.012

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9-18. https://doi.org/10.1016/j.tins.2008.09.012

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools Inc.

Steinborn, M. B., Langner, R., & Huestegge, L. (2017). Mobilizing cognition for speeded action: try-harder instructions promote motivated readiness in the constant-foreperiod paradigm. *Psychological Research*, *81*(6), 1135-1151. https://doi.org/10.1007/s00426-016-0810-1

Welch, R. B., Dutton Hurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Perception & Psychophysics*, *39*(4), 294-300. https://doi.org/10.3758/BF03204939

Supplementary Material.

Table 1. *T* values, corrected p-values (for 23 degrees of freedom) and effect size, as expressed by Cohen's d for each comparison between anisochrony conditions in Experiment 1. The significant differences are highlighted in bold.

Anisochrony (ms)			0			0			0			00			50			00		
			1.43	.48	0.10	1.21	.48	0.10	2.59	.09	0.24	6.90	**10⁻³**	0.47	8.90	**10⁻³**	0.87	13.80	**10⁻³**	1.28
0	.43	.48	.10				0.02	.99	0.00	2.06	.23	0.16	4.86	**10⁻³**	0.40	8.70	**10⁻³**	0.84	13.18	**10⁻³**	1.28
0	.21	.48	.10	.02	.99	0.00				1.99	.23	0.16	4.61	**10⁻³**	0.40	8.16	**10⁻³**	0.83	13.90	**10⁻³**	1.28
0	.59	.09	.24	.06	.23	.16	.99	.23	.16				2.42	.12	0.22	5.19	**10⁻³**	0.61	9.50	**10⁻³**	1.02
00	.9	**10⁻³**	.47	.86	**10⁻³**	.40	.61	**10⁻³**	.40	.42	.12	.22				5.22	**10⁻³**	0.39	9.92	**10⁻³**	0.80
50	.9	**10⁻³**	.87	.70	**10⁻³**	.84	.16	**10⁻³**	.83	.19	**10⁻³**	.61	.22	**10⁻³**	.39				5.11	**10⁻³**	0.43
00	3.8	**10⁻³**	.28	3.18	**10⁻³**	.28	3.90	**10⁻³**	.28	.50	**10⁻³**	.02	.32	**10⁻³**	.80	.11	**10⁻³**	.43			

D. Cutanda, D. Sanabria & A. Correa

Table 2. *T* values, corrected p-values (for 23 degrees of freedom) and effect size, as expressed by Cohen's d for each comparison between anisochrony conditions in Experiment 1. The significant differences are highlighted in bold.

	Anisochrony (ms)																	
	0			**0**			**0**			**00**			**50**			**00**		
	1.19	.00	0.12	0.24	.00	0.02	0.65	.00	0.05	4.07	**10^{-3}**	0.33	5.78	**10^{-3}**	0.52	11.40	**10^{-3}**	0.96
0	.19	.00	.12				.04	.00	.11	.75	.00	.07	2.66	.09	0.23	4.30	**10^{-3}**	0.44	8.36	**10^{-3}**	0.91
0	.24	.00	.02	1.04	.00	0.11				0.38	.00	0.04	3.60	**.01**	0.33	6.07	**10^{-3}**	0.54	12.63	**10^{-3}**	1.01
0	.65	.00	.05	0.75	.00	0.07	.38	.00	.04				3.73	**.01**	0.30	6.23	**10^{-3}**	0.51	10.93	**10^{-3}**	0.98
00	.07	**10^{-3}**	.33	.66	.09	.23	.60	**.01**	.33	.73	**.01**	.30				2.38	.15	0.17	5.96	**10^{-3}**	0.57
50	.78	**10^{-3}**	.52	.30	**10^{-3}**	.44	.07	**10^{-3}**	.54	.23	**10^{-3}**	.51	.38	.15	.17				4.85	**10^{-3}**	0.41
00	1.4	**10^{-3}**	.96	.36	**10^{-3}**	.91	2.6	**10^{-3}**	.01	0.9	**10^{-3}**	.98	.96	**10^{-3}**	.57	.85	**10^{-3}**	.41			

(Left side vertical label: Anisochrony (ms))

Figure S1. Proportion of premature responses as a function of anisochrony (0, 10, 20, 50, 100, 150 and 200 ms) and modality (auditory and visual). Vertical bars represent 95% confidence intervals.
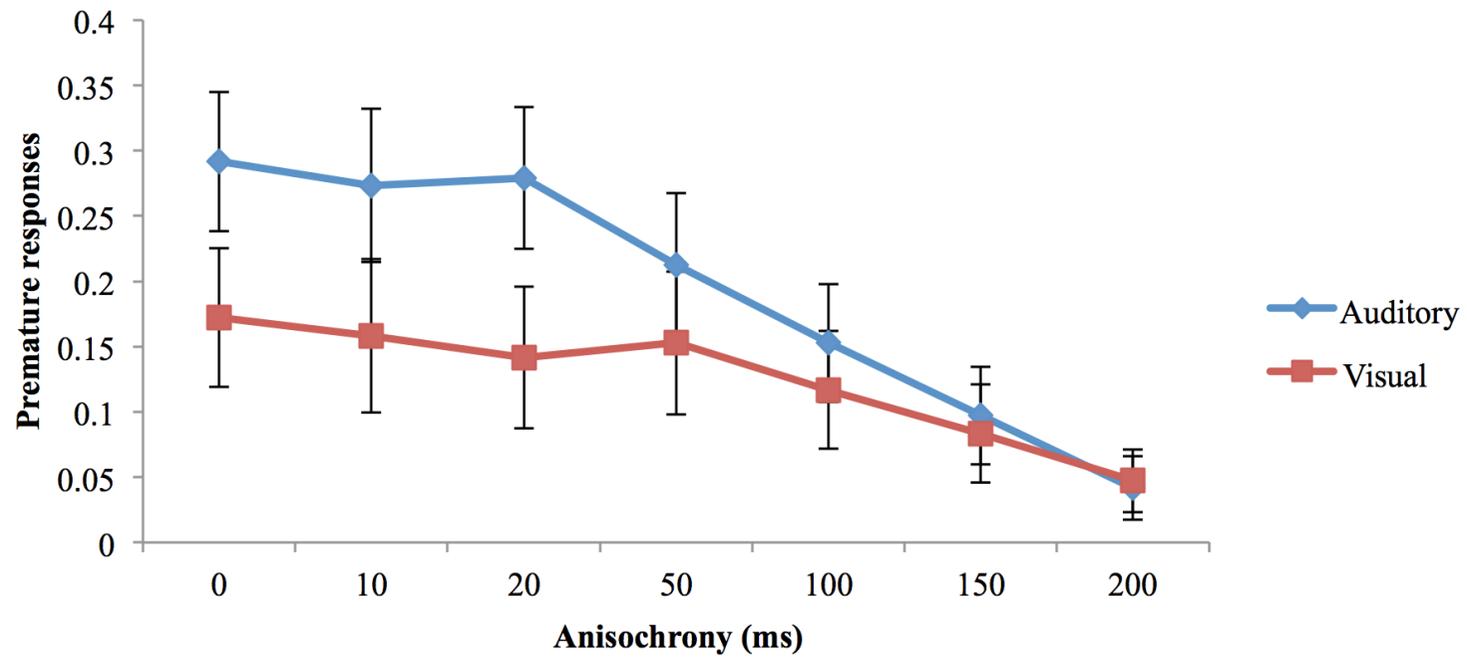
D. Cutanda, D. Sanabria & A. Correa

Figure S2. Proportion of premature responses as a function of anisochrony (0, 10, 20, 50, 100, 150 and 200 ms), modality (auditory and visual) and memory load (0-back and 2-back) in Experiment 2. Vertical bars represent 95% confidence intervals.