Acta Psychologica 142 (2013) 238–244

Contents lists available at SciVerse ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Rhythm implicitly affects temporal orienting of attention across modalities

Deirdre Bolger [a], Wiebke Trost [b], Daniele Schön [a,*]

[a] INS, INSERM & Aix-Marseille University, Marseille, France
[b] Laboratory for Behavioral Neurology and Imaging of Cognition, Department of Neuroscience, Medical School, University of Geneva, Switzerland

ARTICLE INFO

Article history:
Received 18 June 2012
Received in revised form 22 November 2012
Accepted 26 November 2012
Available online xxxx

PsycINFO classification:
2326 Auditory & speech perception
2340 Cognitive processes
2346 Attention

Keywords:
Implicit orienting
Attention
Meter
Temporal structure
Entrainment

ABSTRACT

Here we present two experiments investigating the implicit orienting of attention over time by entrainment to an auditory rhythmic stimulus. In the first experiment, participants carried out a detection and discrimination tasks with auditory and visual targets while listening to an isochronous, auditory sequence, which acted as the entraining stimulus. For the second experiment, we used musical extracts as entraining stimulus, and tested the resulting strength of entrainment with a visual discrimination task. Both experiments used reaction times as a dependent variable. By manipulating the appearance of targets across four selected metrical positions of the auditory entraining stimulus we were able to observe how entraining to a rhythm modulates behavioural responses. That our results were independent of modality gives a new insight into cross-modal interactions between auditory and visual modalities in the context of dynamic attending to auditory temporal structure.

1. Introduction

London (2004) described musical meter as an "emergent property…" that is characterised by multiple, hierarchically-related periodicities over several time scales (Lerdahl & Jackendoff, 1983) and whose basis element is the beat. Thus, beats are best understood as a psychological phenomenon relating to the subjective emphasis of certain events or pulses that are equally spaced in time. It is this phenomenon of beat induction that makes it possible for several people to tap along to a rhythm, all tapping on the same beats and, typically, tapping on those beats that are perceived as salient. Furthermore, given the hierarchical nature of metrical structure in musical rhythms, not all beats are perceived with the same salience, and this leads to the percept of stronger and weaker events within a meter.

The Dynamic Attending Theory (DAT) (Jones & Boltz, 1989; Jones, Boltz, & Kidd, 1982; Large & Jones, 1999) describes these behaviours. DAT focuses on the role of a metrical structure as an active listening strategy, a dynamic structure whose main function is to facilitate future oriented attending, to direct perception and to coordinate behaviour with external events. In conceiving of meter in this way, DAT goes beyond simply modelling meter as a parsing of rhythmic information. The periodicities at each hierarchical level of a metrical structure are described as entraining a listener's neural oscillations and facilitating beat induction (Drake, Penel, & Bigand, 2000; Large, 2000). A computational implementation of DAT (Large & Jones, 1999; Large & Kolen, 1994) models this process as the coupling of internal oscillator periodicities with the external periodicities of the metrical stimulus (entraining stimulus) and predicts the modulation of attentional resources over time in correspondence with the induced percept of strong and weak beats. Consequently, temporal events coinciding with strong beats are more highly anticipated than others and this leads to a focusing of attentional resources around those more anticipated points in time; this has been referred to as anticipatory attending (Jones, 2004).

There is a wealth of research supporting the role of meter in generating temporal expectations and in orienting attention, and thereafter in affecting pitch accuracy judgements (Klein & Jones, 1996; Jones et al., 2002) and temporal just-noticeable differences (Jones & Yee, 1997). The behavioural support for the rhythmic control of attentional orienting in time has been further strengthened by results from neurophysiological studies. Such studies (Snyder & Large, 2005) also provide evidence that an implicit manipulation of expectancy for certain points in time is an innate ability, present among newborn infants (Winkler, Haden, Ladinig, Sziller, & Honing, 2009).

Thus, while there is good support for the DAT, one aspect has, in our opinion, been largely overlooked. Indeed, because the DAT is a rather general theory of attending, it predicts that orienting of attention should not be modality dependent. In other words, attentional

* Corresponding author.
E-mail address: daniele.schon@univ-amu.fr (D. Schön).

0001-6918/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.actpsy.2012.11.012

entrainment should be similar in the auditory and visual modalities, and most importantly for the purpose of the present work, it should be cross-modal. The evidence that attentional entrainment by an auditory meter can also modulate visual attention was only recently presented by Escoffier, Yeo, and Schirmer (2010). In their experiment, subjects performed a speeded response, visual discrimination task in the presence of an auditory four-beat meter; subjects were instructed not to attend to this background meter. The results strongly support a role of auditory meter in attentional modulation by showing that perceptual judgements were faster when the visual stimuli were presented in synchrony with the background auditory meter than when they were presented out of synchrony with it.

Within this context the aim of our investigation was, firstly, to compare the effect of auditory metrical entrainment on both auditory and visual modalities. Secondly, we wanted to study the extent to which entrainment is affected by the difficulty of the task and by changes in attentional demands. Indeed, because dynamic attending is a theory of how attention is directed in time and because attentional resources are limited, it is very possible that dynamic attending will be affected by task attentional demands. For instance, a more demanding task (here discrimination vs detection) may reduce somewhat the effects of a co-occurring entraining stimulus, because less resources can be allocated to attend to it, yielding to a reduction of the amplitude of the attentional oscillations. Thirdly, we wanted to compare the entrainment effect using a simple auditory sequence, such as a metronome, with a more ecological stimulus such as a real musical performance. Finally, rather than using two conditions (on-beat and off-beat) we wanted to have a more "parametric" approach, thus we test entrainment effects at four different metrical positions.

On the basis of these four aims we ran two reaction time experiments investigating the effect of meter on temporal orienting. In the first experiment, participants had to respond as quickly as possible to the presentation of a visual or auditory target in the presence of an isochronous rhythm, which we refer to as an auditory entraining rhythm. The first part of this experiment involved simple detection of target stimuli, while, in the second part, participants had to discriminate targets. In the second experiment the isochronous entraining rhythm was replaced by an excerpt of classical music and participants performed a discrimination task on visual targets. Both experiments were designed such that targets appeared randomly on one of four selected metrical positions (MP) of the auditory entraining rhythm: first beat (MP1), middle beat (MP5) right before the first beat (MP8), and right before the middle beat (MP4). Fig. 1 presents an oscillator model with four metrical levels: the fluctuation of attentional energy is depicted over time as a function of the metrical salience of temporal positions composing our entraining rhythm. Indeed, on the basis of previous models of DAT (Large and Palmer, 2002), it can be considered that the less salient the metrical position is within a given meter, the less attentional energy will be attributed to it. Thus, from this figure one would expect the following order of RTs to targets (ascending, from fastest to slowest): MP1, MP5, MP8, and MP4.

2. Methods

2.1. Experiment 1: isochronous entraining rhythm

2.1.1. Participants

Fourteen participants (mean age: 29, seven females) were recruited. Seven participants had no musical training experience. The other seven participants had between 5 and 30 years of musical practice, a mean of 18 years. Participants were paid 10 Euros.

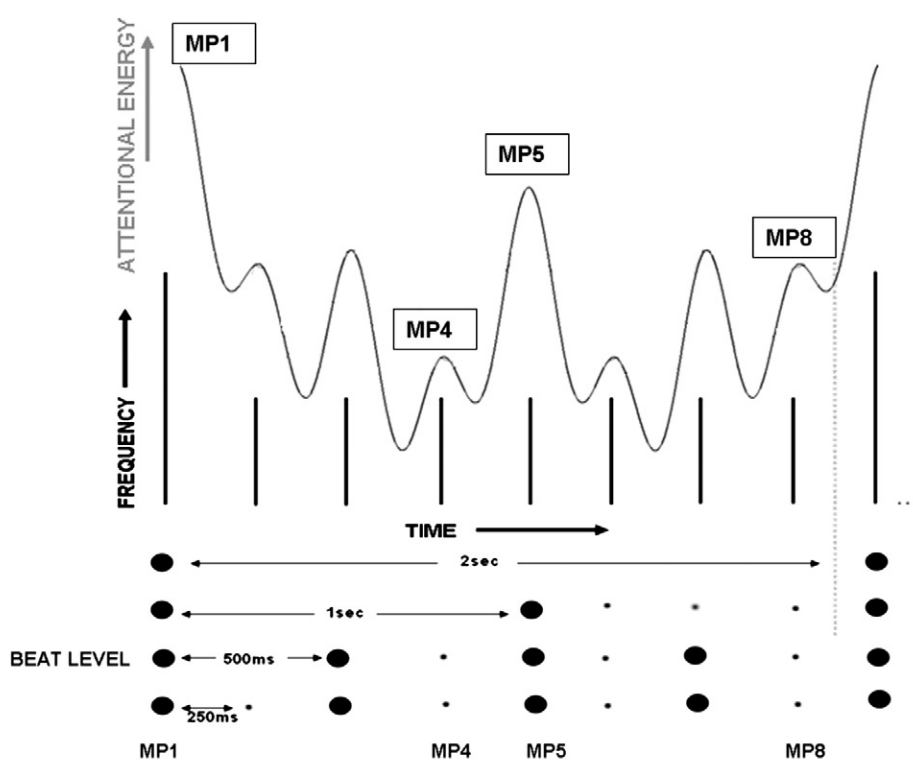

Fig. 1. Metrical structure of a single measure of the 8-tone isochronous entraining rhythm used in experiment 1. The sinusoidal tones composing the 8-tone pattern are represented by the black vertical lines, whose lengths indicate relative pitch level. Four levels of the rhythm's hierarchical, metrical structure are indicated by black circles. The assumed beat level, with an inter-beat interval of 500 ms, is indicated. Assuming this beat level, beat positions are marked by large black circles and non-beat positions by small dots. The four metrical positions (MPs) selected as the independent variables for both experiments are indicated below the tone and beat/non-beat position with which they coincide. The attentional curve, representing the variation of attention level as a function of the metrical structure is shown against the 8-tone measure. Based on the DAT, it represents the total output of a set of four oscillators whose periods are the 250 ms, 500 ms, 1000 ms and 2000 ms. For our purposes it illustrates hypotheses regarding the relative salience, in terms of attentional levels and expectancy, of each of our 4 MPs.

2.1.2. Stimuli

The entraining stimulus was a 5 minute (min) isochronous rhythm composed of repetitions of basic eight-tone sequence, which we will refer to as a measure: sinusoidal tones were used, the first tone of the sequence had a frequency of 880 Hz and the other seven tones had a frequency of 440 Hz. Each tone lasted 100 ms and was presented with an inter-onset interval (IOI) of 250 ms, thus, each eight-tone sequence was 2 s (seconds) in duration. Our auditory entraining stimulus had a simple-duple meter and, given that tapping studies have found that an IOI of 600 ms maximally facilitates beat induction (London, 2004; Moelants, 2002; Parncutt, 1994), we assumed a tactus with a beat period of 500 ms. This was also ascertained during a pilot study in which participants were asked to tap along with the entraining sequence. The same entraining sequence was used for both the detection and discrimination tasks.

For the auditory discrimination task, the targets were a white-noise-like (shaker) sound created using the sound synthesis software, Csound (Barry Vercoe, M.I.T. Labs 1985) and a percussive sound whose tone quality was a hybrid of a wooden and a metallic timbre (Aramaki, Besson, Kronland-Martinet, & Ystad, 2009). The shaker sound was 82 ms in duration and the wood/metal sound was slightly longer with a duration of 150 ms. These sounds were chosen as they contrasted equally well with the sinusoidal tones of the auditory entraining rhythm. For the auditory speeded detection task only the shaker sound was used. The auditory stimuli were delivered at 70 dB via an external sound-card and Sennheiser HD590 headphones.

For the visual discrimination task, the targets, a white "x" and "+", were presented at a viewing angle of 1.4° against a black background at the centre of the screen. The "x" symbol was only used in the visual speeded detection task.

2.1.3. Procedure

2.1.3.1. Detection task. During the detection and discrimination tasks the participants were seated in sound-proofed room. The detection task was composed of two blocks, one for each target modality (auditory and visual) and their order was counter-balanced across participants. Participants carried out a speeded detection task in which they had to react as quickly as possible upon presentation of the target by pressing a response button. For all participants, the response button was held in the right hand. The targets were presented simultaneous to the auditory entraining rhythm, which continued uninterrupted throughout the block. Participants were told that the background sequence did not require their attention. The independent variable manipulated in this experiment was the metrical position of target presentation relative to the background entraining rhythm. Target stimuli could appear at four chosen temporal positions: on the first tone of the sequence (metrical position 1 or MP1), on the fourth tone (metrical position 4 or MP4), on the fifth tone (metrical position 5 or MP5) and on the eighth tone (metrical position 8 or MP8) (Fig. 1). Assuming a beat level with a 500 ms period, MP1 and MP5 correspond to the first and third beats, respectively. MP4 and MP8 correspond to the off-beat positions preceding MP5 and MP1. The first presentation of a target within a session could not occur before two repetitions of the basic eight-tone pattern, i.e. earlier than 4 s, and, while the four metrical positions were probed randomly, the same MP could not be probed consecutively any more than three times so as to limit any possible learning effect that might facilitate RTs. Furthermore, the minimum and maximum permitted duration between target presentation was 2 s and 8 s, respectively, thereby avoiding the probing of two MPs within a single measure. We used a 'non-ageing' distribution in order to keep a posteriori probability for the stimulus constant as time passes. This was done in order to the render stimulus onset asynchrony (SOA) and the RTs independent (Näätänen, 1971). Each block (auditory and visual) lasted approximately 16 min and was subdivided into three mini blocks of 5 min 30 s duration. The random probing of temporal positions by target presentation differed in each block.

2.1.3.2. Discrimination task. The design of the discrimination task was the same as that for the detection task described above, with the exception that, instead of a single target stimulus, two targets of the same modality were used. Participants held a response button in each hand and were informed as to the response side (right or left) associated with each of the two target stimuli. Association of a target with the button in the left or right hand was counterbalanced across participants. The task, therefore, was to discriminate between the two target stimuli as quickly and as accurately as possible in response to target presentation by pressing the correct response button (left or right) as promptly as possible. Target modality was the same for each session and each of the two targets appeared with equal probability ($p = 0.5$). The order of blocks (auditory or visual targets) was counterbalanced across participants.

In both detection and discrimination tasks and for both modalities, participants were presented with 54 trials/metrical position (e.g. visual detection task: 216 trials). The distribution of the targetSOA for each condition was the same to control for temporal serial order effects. The order of tasks (detection and discrimination) was counterbalanced across subjects.

The experiment was carried out using Presentation software (Neurobehavioral Systems) for the delivery of stimuli and experimental control. Participants were seated comfortably in a silent room in front of a screen, at a distance of 60 cm. The experiment lasted 1 h in total.

A pilot study was conducted to verify if participants could distinguish target appearance at different metrical positions (e.g. MP1 and MP8). Specifically, the aim was to investigate if there was a migration effect between positions close in time (e.g. MP8 to MP1). We used the same experimental design and stimuli employed in the visual and auditory detection tasks described above. Eight participants took part in the study, four of whom were musicians. They were instructed to indicate if targets were presented on or off the beat. The study was composed of twenty trials and participants responded correctly on between 90 and 100% of trials.

2.1.4. Analysis and results

A $2 \times 2 \times 4$, repeated measures ANOVA was carried out for factors Task, Modality and Meter on the RT results for the detection and discrimination tasks. This revealed a main effect of the variable Task ($F(1,13) = 212.21$, $p < 0.0001$) with slower RTs in the discrimination than the detection task. Most importantly, the main effects of Meter ($F(3,39) = 41.19$, partial $\eta^2 = 0.75$, $p < 0.0001$) and of the Task × Meter interaction ($F(3,39) = 12.21$, $p < 0.0001$) were significant. The Task × Meter × Modality interaction ($F(3,39) = 0.42$, $p = 0.74$) was not significant.

Post-hoc analyses showed that in both visual and auditory detection tasks RTs to targets presented at MP1 and MP8 were highly similar and significantly faster than RTs to targets presented at MP4 and MP5 (always $p < 0.005$, see Fig. 2). RTs to targets at MP4 and MP5 did not show significant differences. In the discrimination task, targets at MP4 yielded RTs that were significantly slower compared to MP1, MP5 and MP8, ($p < 0.005$). RTs to MP1, MP5 and MP8 targets did not differ significantly from each other.

There was no significant effect of the variable Modality ($F(1,13) = 1.25$, $p = 0.28$), nor of the interaction Modality × Meter ($F(3,39) = 0.38$, $p = 0.77$).

Finally, ANOVA on accuracy measures in the discrimination task did not show a significant effect of meter on accuracy, for both modalities ($F(3,39) = 0.47$, $p = 0.7$). The percentage of correct responses for the auditory and visual discrimination tasks was 95%.

2.2. Experiment II: musical stimulus

In the first experiment we observed a difference between the meter effect in the detection and discrimination tasks such that, in the discrimination task, RTs to targets presented on MP1, MP5 and

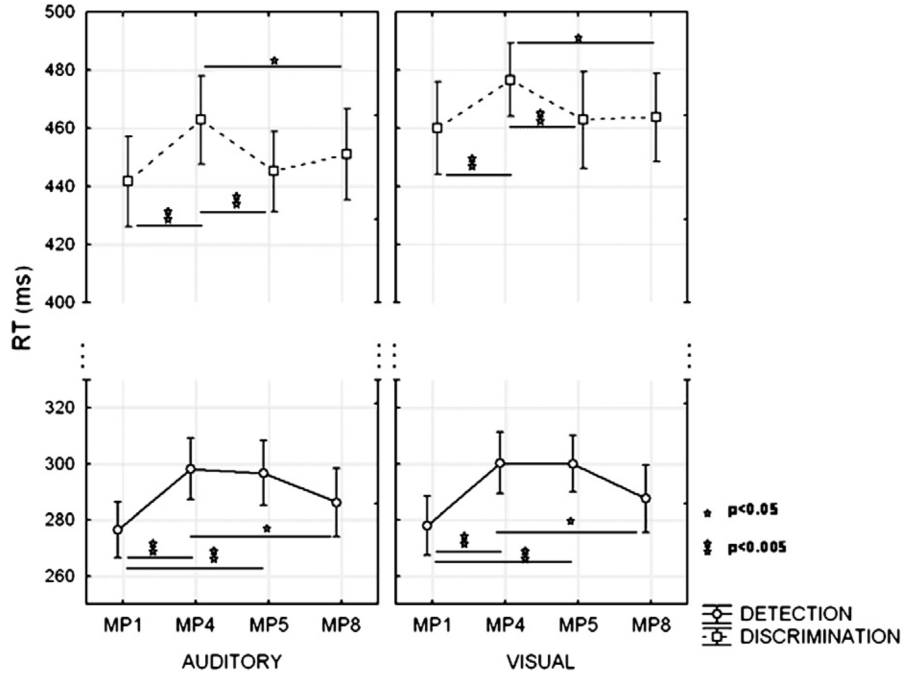

Fig. 2. Results of experiment 1, showing the mean RTs of the four metrical positions (MPs) across all participants for each task and each modality (detection and discrimination). Significant RT differences between metrical positions are indicated by a horizontal line. The confidence intervals were computed as described in Loftus and Masson (1994) and take into account inter-subject variability.

MP8 were similar, with only those relating to MP4 showing a significant difference. We hypothesised that this may be due to the increased difficulty of the discrimination task, compared to the detection task and decided that one solution to this may be to use more ecological stimuli to "boost" entrainment. Indeed, it has been shown that, compared to a mechanical performance (i.e. metronometrical), real performances facilitate a full access to the metrical structure (Drake et al., 2000). Moreover, the isochronous rhythm had a higher pitch falling on the strong beat (MP1) of each measure, and although we do not believe that this pitch accent is driving the results of experiment 1, we felt that carrying out the same task using an auditory stimulus with greater variation in pitch as well as rhythm would be important.

Therefore, in this second experiment participants performed a visual discrimination task using excerpts of real recordings of classical music as the background entraining stimulus. Our main reason for carrying out this task for the visual modality only was that the discriminability/saliency of auditory targets was found to vary at different points of the musical excerpt, with the implication that extracting an overall effect of meter in the auditory discrimination task would not be possible.

2.2.1. Participants

Twenty-two participants (mean age: 31, 10 females) were recruited. Six participants had no musical training experience. The other participants had between 7 and 25 years of musical practice, thus a mean of 19 years musical practice (the difference in musical training experience between Exp 1 and 2 was not significant, $p = 0.12$). Participants were paid 10 Euros.

2.2.2. Stimuli

In this second experiment, in place of the isochronous entraining meter we employed recordings of real musical performances. Excerpts from four pieces of western classical music with a clear tactus and metrical structure were selected. In addition, three other criteria were applied when selecting the musical excerpts, all to ensure ease of metrical entrainment among participants. Firstly, all selected musical excerpts had a clear binary metrical structure. Secondly, only excerpts containing many notes per measure were chosen in order to facilitate meter perception. Thirdly, the musical excerpts had an average tactus of around 600 ms. Lastly, pieces with a maximum of two instruments were selected so as to reduce the complexity of the acoustical signal of the musical stimulus. As part of the selection of the musical excerpts, a number of pieces that differed in terms of style and instrumentation and which adhered to the above-mentioned criteria were chosen. Several excerpts were presented to six non-musicians in a pilot experiment and the participants were asked to tap along with the excerpts and to rate the difficulty in detecting the meter of each. On the basis of this study the four pieces which emerged as easiest in terms of metrical structure were chosen. Participants listened to all pieces in a random order

- Johann Sebastian Bach: Partita for solo violin No. 2, BWV 1003, Allegro
- Wolfgang Amadeus Mozart: Piano Sonata in C major, KV 330, Allegro
- Gioachino Antonio Rossini: Transcription for clarinet and piano
- Antonio Vivaldi: Sonata No. 6 for flute and harpsichord from "Il Pastor Fido", Allegro.

2.2.3. Experimental design

2.2.3.1. Procedure. The task and design constraints were identical to those used in the discrimination visual task with the isochronous sequence. 53 visual targets were presented in each of the four experimental conditions.

The exact latency of the time points corresponding to the different conditions was detected using the software, Adobe Audition. The onsets were defined semi-automatically: using a spectral representation of the music a criterion of a rise of 2 dB in a window of 5 ms was firstly applied and then the results were adjusted manually if necessary. The musical instances selected to present visual stimuli were chosen in

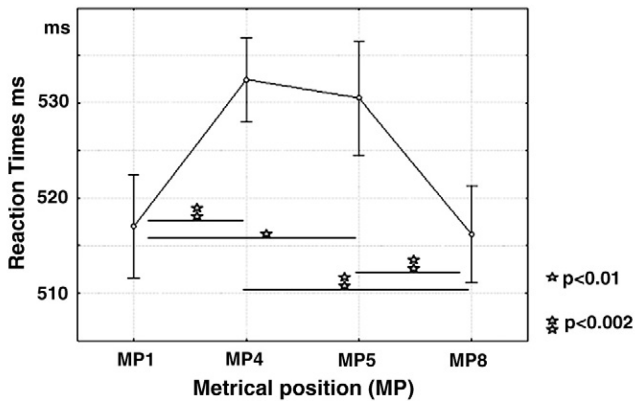

Fig. 3. Result of experiment 2 showing mean RTs for each of the four MPs across all participants for the visual discrimination task using musical excerpts as the entraining stimuli. Significance differences in RTs between metrical positions are indicated by horizontal lines, with the stars representing the level of significance. The confidence intervals were computed on the basis of within-subject variance (Loftus & Masson, 1994).

such a way that they did not differ significantly across conditions in terms of amplitude of the signal (ANOVA on the root mean square (RMS) values of the four conditions, $p>0.99$). This was done to prevent potential differences in loudness that may render the interpretation of results difficult. The global mean volume (RMS) was normalised across the four different pieces.

2.2.4. Analysis and results

One-way repeated measures ANOVA was carried out for the variable Meter on the RT results. The main effect of Meter was significant ($F(3,63)=8.7$, partial $\eta^2=0.55$, $p<0.001$). Fig. 3 presents the mean RT results for all participants as a function of the four metrical positions.

Post-hoc analyses showed that RTs to targets presented at MP1 (517 ms) and MP8 (516) were significantly faster than RTs to target positions MP4 (532 ms, p always <0.002) and MP5 (530 ms p always <0.01). RTs to targets at MP4 and MP5 ($p>0.9$), did not show significant differences, nor RTs to targets at MP1 and MP8 ($p>0.9$).

3. Discussion

Our aims were to compare the effect of auditory metrical entrainment in both auditory and visual modalities and to study the extent to which the entrainment effect would be affected by the difficulty of the task and by the ecological validity of the entrainment stimulus. On the basis of these aims we ran two separate experiments investigating the effect of meter on temporal attentional orienting by probing four different metrical positions. In the first experiment, we used a speeded detection and a discrimination task wherein participants had to respond to the presentation of visual or auditory targets in the presence of an isochronous auditory rhythm; on each trial, target presentation coinciding with any one of the four selected metrical positions. In the second experiment participants performed only a visual discrimination task while listening to several excerpts of classical music, which served as the entraining stimulus.

Our results showed that, in both the speeded detection task and the discrimination task, the observed effect of meter on performance was independent of the modality of the target. This suggests that the meter-driven orienting of attention over time affected the processing of visual and auditory targets equally. Moreover, that visual target processing is affected by auditory entrainment was also confirmed using more ecological stimuli such as recordings of classical music performances. This is an interesting observation in terms of the cross-modal links between auditory temporal structure and visual processing. Indeed, fields other than entrainment have found evidence of cross-modal links between auditory and visual modalities (Lange & Röder, 2006; Shomstein & Yantis, 2004; Spence & Driver, 1996; Vroomen & de Gelder, 2000). Our results confirm and extend the findings of Escoffier et al. (2010), by showing that not only is visual perceptual judgement affected by auditory entrainment, but perception of visual and auditory targets is also influenced in a similar manner.

While our results allow us to address cross-modal attentional effects, we accept that it is not possible to draw definite conclusions concerning the influence of meter entrainment on visual processing based on this behavioural study alone. This is especially true of the detection task but is also the case for the discrimination task; in both it is difficult to disentangle perceptual processing from motor preparation on the basis of RT data alone (Correa, Lupianez, Milliken, & Tudela, 2004; Correa, Lupiáñez, & Tudela, 2005; Nobre, 2001). Moreover, in the discrimination task results, the number of incorrect responses is not sufficient to offer clear evidence of a contribution of perceptual preparation. Finally, the discrimination response was given under a time pressure and so the influence of motor preparation cannot be discounted (Correa et al., 2005). However, previous studies have provided evidence of a facilitation of early perceptual processing by explicit temporal cues in both auditory and visual modalities (Correa, Lupiáñez, & Tudela, 2006; Lange, Rösler, & Röder, 2003), but these employed tasks that were perceptually demanding and whose interpretation relied upon electrophysiological correlates. Still, while accepting that, we point to the fact that the performance differences between the strongest and the weakest metrical position are consistent and cannot be ignored. Furthermore, the fact that this effect is preserved across tasks is an encouraging indication of a facilitation of perceptual and cross-modal processing by temporal attention. Of course, this hypothesis would need to be confirmed by further work using either electrophysiological correlates of perceptual processing or psychophysical methods (e.g. just-noticeable-differences, Correa et al., 2006).

In our first experiment, in which the entraining stimulus was an isochronous rhythm, our results showed a significant difference in the effect of meter on performance between the detection task and discrimination task. While in the detection task RTs to MP4 and MP5 were slower than MP1 and MP8, in the discrimination task the effect was such that only RTs to MP4 targets differ significantly from the other metrical positions (Fig. 2). We suggest that differences in attentional demands between detection and discrimination tasks underlie this difference. Indeed, while the effects of the task difficulty on the dynamics of attention, as predicted by DAT (Fig. 1), are certainly not linear, one may expect that the differences predicted by the model as a function of metrical levels may be globally reduced by a co-occurring demanding task. Nonetheless, the literature is not clear cut concerning the extent to which executive functions or more generally top-down mechanisms may affect dynamic attending. For instance, while Schwartze, Rothermich, Schmidt-Kassow, and Kotz (2011) consider that temporal regularity is affected by top-down attentional processing, others argue that temporal preparation driven by rhythms is mostly stimulus driven and does not require (and is not affected by) executive functions (De la Rosa, Sanabria, Capizzi, & Correa, 2012).

Interestingly, the results of the visual discrimination experiment using a musical entraining stimulus were very similar to the results found in the detection task using the isochronous stimulus. This is important for several reasons. Firstly, it shows that the cross-modal effect of meter shown here is not simply restricted to isochronous stimuli, but is a general property of structured rhythms that also applies to highly variable and more musical rhythmic patterns.

Secondly, it shows that the use of more ecological stimuli is not only possible, but also advisable, insofar as it seems to magnify the level of entrainment. While the isochronous rhythm only allows differentiating discrimination of visual targets presented at the weakest position from the three other positions, the use of music allows differentiating discrimination of visual targets presented at the two weaker positions from the two stronger positions. This possibly takes place via a higher and more stable effect of arousal driven by real music compared to more standard laboratory stimuli. Indeed, such an event is predicted by DAT; a more complex rhythm engages a greater number of oscillators whose coupling yields more stable beat periods (Patel, Iversen, Chen, & Repp, 2005). Also the temporal variations intrinsic to each real performance facilitate a full access to the metrical structure (Drake et al., 2000). Also one should consider that the music recording we used was possibly not the most entraining musical excerpts and pop, traditional or jazz music might be more rhythmically entraining than Bach and Vivaldi.

Finally, given that in Experiment 1 the tone at the strongest metrical position (MP1) also had the highest pitch of each group of eight tones, one could argue that it is very hard to know if the results relating to MP1 are due to an acoustic saliency change or to the meter. Other type of rhythmical structures could be possibly used to better control for this type of bias (Escoffier et al., 2010, Winkler et al., 2009). Nonetheless, when using music as an entraining stimulus, we obtained similar results. Because, in music, pitch is not a consistent marker of metrical units, the similarity of the results obtained using isochronous and musical entrainment stimuli strongly favours the interpretation in terms of meter in both experiments. Moreover, in Experiment 1, the acoustic saliency between positions 1 and 8 was different but RTs were similar, and while the acoustic saliency between positions 8, 4 and 5 were identical, their RT differed (at least in the detection task).

Overall, our results strongly suggest that the metrical structure of an entraining rhythm does indeed implicitly modulate the directing of attentional resources in time, with an associated facilitation of task performance (faster RTs). This confirms previous findings in implicit (Ellis & Jones, 2010; Praamstra, Kourtis, Kwok, & Oostenveld, 2006) and also in explicit (Barnes & Jones, 2000; Coull & Nobre, 1998; Miniussi, Wilding, Coull, & Nobre, 1999; Nobre, 2001) cueing paradigms, which show faster RTs to validly cued targets than to invalidly cued targets.

Nonetheless, our results stray from the predictions illustrated in Fig. 1. According to the model predictions, one would expect faster RTs at MP5 compared to both MP4 and MP8. However, our results consistently show faster RTs to targets at MP8 than to those at MP5 in the speeded detection task and in the discrimination task using musical excerpts. Moreover, in all tasks there is a consistent similarity between RTs to MP1 and MP8 targets. Therefore, when considering the two different hypotheses described above, it seems that targets presented at MP8 are "attracted" by MP1. We emphasise that this is not the same as saying that targets presented at MP8 are perceived as if presented at MP1. Indeed, a pilot experiment investigating this issue showed that participants differentiated these two positions with high accuracy, thus lending support to the hypothesis that RTs to targets at MP8 and MP1 were similar because temporal expectancy for them was similar. Furthermore, a similar relationship was observed between MP4 and MP5 RTs: though RTs to targets presented at MP4 were consistently slower than those to MP5 this difference did not reach significance. Again, this suggests that temporal preparation for these MPs was also similar. Thus, it seems that events falling within a given temporal window around the first beat (in our case 250 ms before) benefit from its increased temporal expectancy. This is also in line with recent findings showing that rhythmic entrainment effects are somewhat anticipatory and can be found in intervals that are earlier than the interval matching the entrainment rhythm (Sanabria et al., 2011). This yields a scenario wherein MP1 and MP8 play the role of highly expected positions while MP4 and MP5 constitute less expected positions. This can be accounted for by a two-oscillator model with a periodicity of one measure and half measure (see Fig. 4). Conceiving of the distribution of attentional energy over time on the basis of such a model presents a situation in which a greater level of attentional energy is directed towards events falling in the vicinity of MP1 than towards events that occur proximate to MP5. This is depicted in Fig. 4, with MP8 having a greater level of attentional energy than MP5 despite being an off-beat position.

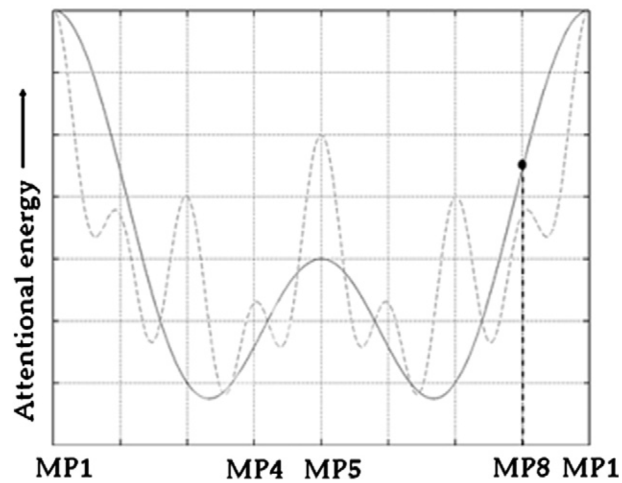

Fig. 4. Illustration of "attractor" hypothesis regarding the similarity of behavioural results observed for MP8 (an off-beat position) and MP1 (strong beat position). It shows the output of a two oscillator model with periodicities of one measure (2000 ms) and a half a measure (1000 ms). The position of MP8 on the attention curve is marked highlighting that given such a model, the attentional energy attributed to MP8 tends towards that of MP1 and exceeds that of MP5.

To summarise, in our experiments we have shown an effect of metrical entrainment on responses to both visual and auditory targets. While our RT paradigm cannot fully address the question of whether metrical entrainment can enhance stimulus processing, the robustness of our results suggests that this is worth further investigation by psychophysiological studies. Our findings build on existing research on implicit attentional orienting by showing that metrical entrainment by both simple isochronous meter and complex musical stimuli can affect orienting of attentional resources in a purely implicit manner over time facilitating (or otherwise), as a result, responses to stimuli presented at different points in time, and, doing so independently of stimulus modality.

Acknowledgements

The series of experiments reported above were conducted thanks to the support of the French National Agency (ANR-07-NEURO-033-01). We thank Dr. Jennifer Coull and two anonymous reviewers for helpful comments on a previous version of this manuscript.

References

Aramaki, M., Besson, M., Kronland-Martinet, R., & Ystad, S. (2009). Timbre perception of sounds from impacted materials: behavioral, electrophysiological and acoustic approaches. Computer modeling and retrieval. Genesis of meaning in sound and music. *Series lecture notes in computer science, 5493/2009* (pp. 1–17). Berlin/Heidelberg: Springer.

Barnes, R., & Jones, M. R. (2000). Expectancy, attention and time. *Cognitive Psychology*, *41*, 254–311.

Correa, Á., Lupianez, J., Milliken, B., & Tudela, P. (2004). Endogenous temporal orienting of attention in detection and discrimination tasks. *Perception & Psychophysics*, *66*(2), 264–278.

Correa, Á., Lupiáñez, J., & Tudela, P. (2005). Attentional preparation based on temporal expectancy modulates processing at the perceptual level. *Psychonomic Bulletin and Review*, *12*, 328–334.

Correa, Á., Lupiáñez, J., & Tudela, P. (2006). The attentional mechanism of temporal orienting: determinants and attributes. *Experimental Brain Research*, *169*, 58–68.

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: The neural systems for directing attention to spatial locations and to time intervals as revealed by both pet and fMRI. *Neuroscience*, *18*, 7426–7435.

De la Rosa, M. D., Sanabria, D., Capizzi, M., & Correa, A. (2012). Temporal preparation driven by rhythms is resistant to working memory interference. *Frontiers in Psychology*, *3*, 308.

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception*, *18*(1), 1–23.

Ellis, R. J., & Jones, M. R. (2010). Temporal context and choice reaction time. *Attention, Perception, & Psychophysics*, *72*, 2274–2288.

Escoffier, N., Yeo, D., & Schirmer, A. (2010). Unattended musical beats enhance visual processing. *Acta Psychologica*, *135*(1), 12–16.

Jones, M. R. (2004). Attention and timing. In J. G. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 49–85). New York: Academic Press.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459–491.

Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics*, *32*, 211–218.

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.

Jones, M. R., & Yee, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology. Human Perception and Performance*, *23*, 693–709.

Klein, J. M., & Jones, M. R. (1996). Effects of attentional set and rhythmic complexity on attending. *Perception & Psychophysics*, *58*(1), 34–46.

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*, 1–37.

Lange, K., & Röder, B. (2006). Orienting attention to points in time improves processing both within and across modalities. *Journal of Cognitive Neuroscience*, *18*, 715–729.

Lange, K., Rösler, F., & Röder, B. (2003). Early processing stages are modulated when auditory stimuli are presented at an attended moment in time: An event-related potential study. *Psychophysiology*, *40*, 806–817.

Large, E. W. (2000). On synchronizing movements to music. *Human Movement Science*, *19*, 527–566.

Large, E., & Jones, M. R. (1999). The dynamics of attending: How we track time-varying events. *Psychological Review*, *106*, 119–159.

Large, E., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, *6*, 177–208.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge MA: MIT Press.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review*, *1*, 476–490.

London, J. (2004). *Hearing in time.* Oxford University Press.

Miniussi, C., Wilding, E. L., Coull, J. T., & Nobre, A. C. (1999). Orienting attention in time: Modulation of brain potentials. *Brain*, *122*, 1507–1518.

Moelants, D. (2002). Preferred tempo reconsidered. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, 2002* (pp. 580–583). Adelaide: Causal Productions.

Näätänen, R. (1971). Non-aging foreperiods and simple reaction time. *Acta Psychologica*, *35*, 316–327.

Nobre, A. C. (2001). Orienting attention to instants in time. *Neuropsychologia*, *39*, 1317–1328.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*, 409–464.

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, *163*, 226–238.

Praamstra, P., Kourtis, D., Kwok, H. F., & Oostenveld, R. (2006). Neurophysiology of implicit timing in serial choice reaction-time performance. *The Journal of Neuroscience*, *26*(20), 5448–5455.

Sanabria, D., Capizzi, M., & Correa, A. (2011). Rhythms that speed you up. *Journal of Experimental Psychology. Human Perception and Performance*, *37*, 236–244.

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biological Psychology*, *87*, 146–151.

Shomstein, S., & Yantis, S. (2004). Control of attention shifts between vision and audition in human cortex. *The Journal of Neuroscience*, *24*(47), 10702–10706.

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, *24*, 117–126.

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology. Human Perception and Performance*, *22*(4), 1005–1030.

Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organisation on vision. *Journal of Experimental Psychology. Human Perception and Performance*, *26*(5), 1583–1590.

Winkler, I., Haden, G. H., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences (PNAS)*, *106*(7), 2468–2471.