There's something about that groove: Rhythm improves detection of audio but not vibrotactile asynchronies

Andrew P. Lauzon*[1,2], Frank A. Russo†[2], and Laurence R. Harris‡[1]

[1] Department of Psychology, York University, 4700 Keele St., Toronto, ON, Canada, M3J 1P3.

[2] Department of Psychology, Ryerson University, 350 Victoria St., Toronto, ON, Canada, M5B 2K3.

1 Introduction

It is well established that context can powerfully influence sensory perception and that rhythm can induce entrainment in the form of regularly occurring evoked potentials in neural activity. These evoked potentials are defined by the interval between salient beats, and are thus oscillatory in nature. These neural oscillations are presumed to modulate (a) sensory input responses (Schroeder & Lakatos, 2009) as well as (b) attention, an idea described by Jones, Moynihan, MacKenzie and Puente (2002) to explain their finding that judgments regarding small pitch deviations were most accurate when target stimuli occurred closest to an expected temporal interval (i.e. on the beat).

The present study examined the influence of an isochronous rhythm on judgements of simultaneity in auditory and vibrotactile modalities by measuring thresholds of detection (TOD) for tones presented with slight asynchronies, both with and without the context of an isochronous pattern. A significant improvement in detection when accompanied by isochrony would provide support for the previously mentioned theoretical framework, and a difference in thresholds obtained between the auditory and vibrotactile modalities would be in line with the notion of an auditory advantage for temporal processing.

2 Method

2.1 Participants

A convenience sample of 10 volunteers was recruited by word-of-mouth and consisted of four females and six males (age range 22-48; mean age = 31.7); all had normal hearing, and musical experience ranged from professional musicians to those with no musical training. Hearing and musical experience were determined by a questionnaire which was completed prior to the experiment. Participants were not paid. Experiments were approved by the York University and Ryerson Research Ethics Board and adhered to the Treaty of Helsinki.

2.2 Apparatus

Experimental trials were run using Cycling '74 Max MSP software on a 2010 Macbook Pro with an RME Fireface 400 firewire audio interface. Audio for the auditory conditions was delivered to participants through Sennheiser HD518 over-ear headphones. Vibrotactile stimuli also originated as audio, and were sent to vibrating voice coils embedded in the seat and back of a slightly reclining padded chair (Emoti-chair; Karam, Russo, & Fels, 2009) in which participants would sit for the duration of the experiment. A pre-study was conducted in order to equalize perceived magnitude between the auditory and vibrotactile stimuli.

Tactaid VBW32 transducers were attached to the participants' mastoid bones to deliver pink noise for the purpose of masking auditory detection of bone conducted vibrations potentially received via the voice coils. Pink noise was also delivered through the headphones to mask any additional airborne sound that was emitted by the voice coils in the chair.

2.3 Stimuli

Target stimuli consisted of a pair of pure tones (200 Hz and 300 Hz) presented either in perfect synchrony or with one of 10 possible stimulus onset asynchronies (SOA). The same tones were used for both modalities, however SOAs for the auditory conditions ranged from 5-23 milliseconds (ms) whereas vibrotactile SOAs ranged from 10-190 ms. Target stimuli were presented in one of the three different rhythmic contexts for both modalities, for a total of six experimental conditions:

Regular Rhythm (RR) : Eight beats occurred with an inter stimulus interval (ISI) of 500 ms, or 120 beats per minute (BPM). Beats 1-6 and 8 were the context stimuli, which consisted of the 200 Hz and 300 Hz pure tones played in perfect synchrony and beat 7 was the target stimulus which was the same two tones, either identical to the context stimuli or with a slight SOA.

Irregular Rhythm (IR) : Target stimuli were present within the context of an irregularly occurring, unpredictable beat sequence. This was identical to the regular rhythm condition, with the exception that the ISI duration changed on every beat for the first six beats. Target stimuli occurred at the same relative time point within each trial, as compared to all other conditions.

No Rhythm (NR) : Target stimuli were presented unaccompanied (with no context). This condition was identical to the regular rhythm, except the context stimuli (beats 1-6 and 8) were silent.

*alauzon@my.yorku.ca

†russo@ryerson.ca

‡harris@yorku.ca

Figure 1: Graphical representation of the three rhythmic context conditions with asynchronous target stimuli.

A two-alternative forced choice procedure was used. For each trial, participants would hear or feel the target stimulus twice: once with the asynchrony at beat 7 and once with no asynchrony, to decide which of the two presentations contained the SOA, the order of which was randomized. Answers were indicated by entering either "1" or "2" on a computer keyboard. A block design was employed with each block consisting of one of the three rhythmic contexts, the order of which was counterbalanced.

2.4 Data Analysis

Data were plotted as a function of SOA value, and for each participant a percentage correct score was calculated for each SOA. A logistic curve was fit to each participant's data, and each participant's TOD was determined as the SOA value at which they achieved 75% correct. Standard deviation for each participant was extracted from the slope of the curve, and overall variability was compared between modalities as a measure of response consistency.

Separate repeated-measures ANOVAs were performed for each modality to compare TODs, and a two-way repeated measures ANOVA was performed to compare overall variability between the two modalities. Pairwise comparisons used the Bonferroni correction.

3 Results

3.1 Detection Thresholds

A significant effect of rhythmic context was found in the auditory conditions, $F(2, 18) = 3.56$, $p = .05$, partial $\eta^2 = 0.283$ (large). Pairwise comparisons revealed a significant difference between the no rhythm ($M = 8.97$, $SD = 3.06$) and regular rhythm ($M = 6.12$, $SD = 1.60$) conditions which corresponds to a ~32% reduction in thresholds ($MD = 2.85$, $SE = .929$, $p = .04$), but no difference when comparing either of the conditions to irregular rhythm ($M = 7.66$, $SD = 2.76$). A repeated-measures ANOVA performed on the vibrotactile conditions failed to reach significance $F(2, 18) = 2.86$, $p = .084$, partial $\eta^2 = 0.241$.

3.2 Variability

A two-way repeated measures ANOVA compared variability between all conditions (both modalities and all rhythmic contexts) and showed a highly significant difference $F(1, 9) = 23.34$, $p = .001$, partial $\eta^2 = .722$. A pairwise comparison of the overall mean variability scores of each modality ($MD = 6.23$, $SE = 1.29$, $p = .001$) was also significant. This suggests the vibrotactile conditions were more difficult as a whole.

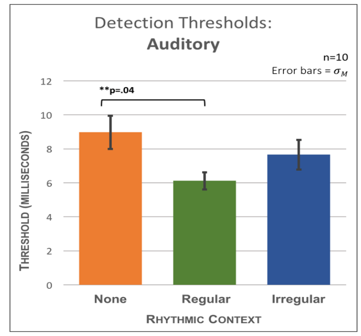

Figure 2: Detection thresholds for the auditory conditions. Note the significant reduction (p = .04) in the regular rhythm condition compared to no rhythm. No significant differences were found between any of the vibrotactile conditions.

4 Discussion and Conclusions

These results support the hypothesis that context in the form of an isochronous rhythmic pattern does indeed affect simultaneity judgements in the auditory domain. It also contributes further support to the hypothesis of privileged temporal processing for auditory stimuli and prompts questions about vibrotactile perception, cross-modal plasticity, and higher order processing of temporal information.

The high variability and lack of significant results in the vibrotactile conditions suggests a failure to entrain, and may simply be due to a lack of acuity in lower-level somatosensory processes, which are unsuccessful in transmitting neural impulses with enough temporal precision to synchronize attentional dynamics, expectations, or neural oscillations. On the other hand, an experience-dependent explanation could support a more top-down, cortically mediated process in which timing information received via somatosensory receptors is robust, but processed inefficiently and subject to improvement with practice.

References

Karam, M., Russo, F. A., & Fels, D. I. (2009). Designing the Model Human Cochlea: An Ambient Crossmodal Audio-Tactile Display. *IEEE Transactions on Haptics*, *2*(3), 160–169.

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.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18.