

Bouncing the Network: A Dynamical Systems Model of Auditory-Vestibular Interactions Underlying Infants' Perception of Musical Rhythm

Parker Tichko¹, Ji Chul Kim² & Ed Large^{2,3}

Department of Music, Northeastern University¹

Department of Psychological Sciences, University of Connecticut²

Department of Physics, University of Connecticut³

A preprint is available here!

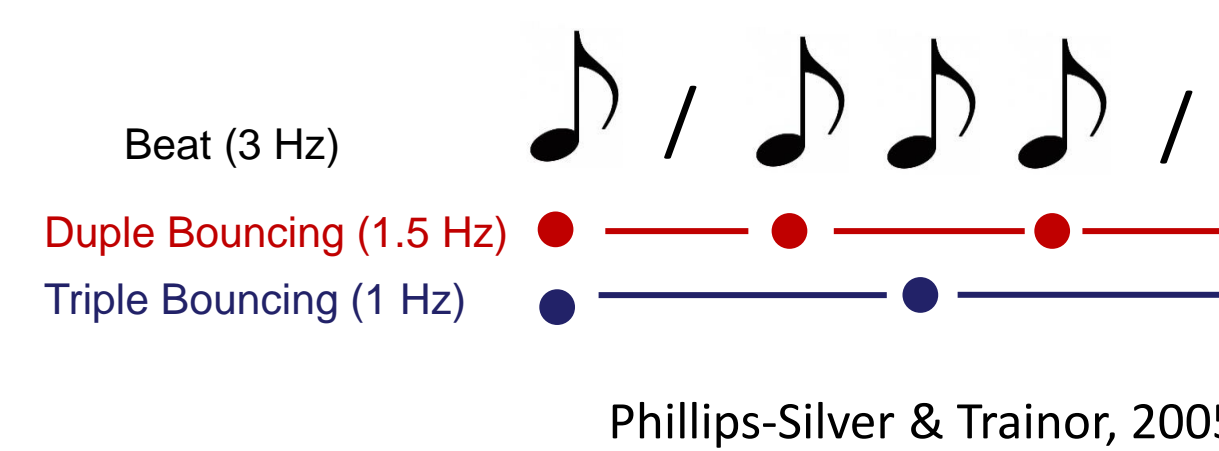
<https://bit.ly/2Ygu7Jg>



UCONN | UNIVERSITY OF CONNECTICUT

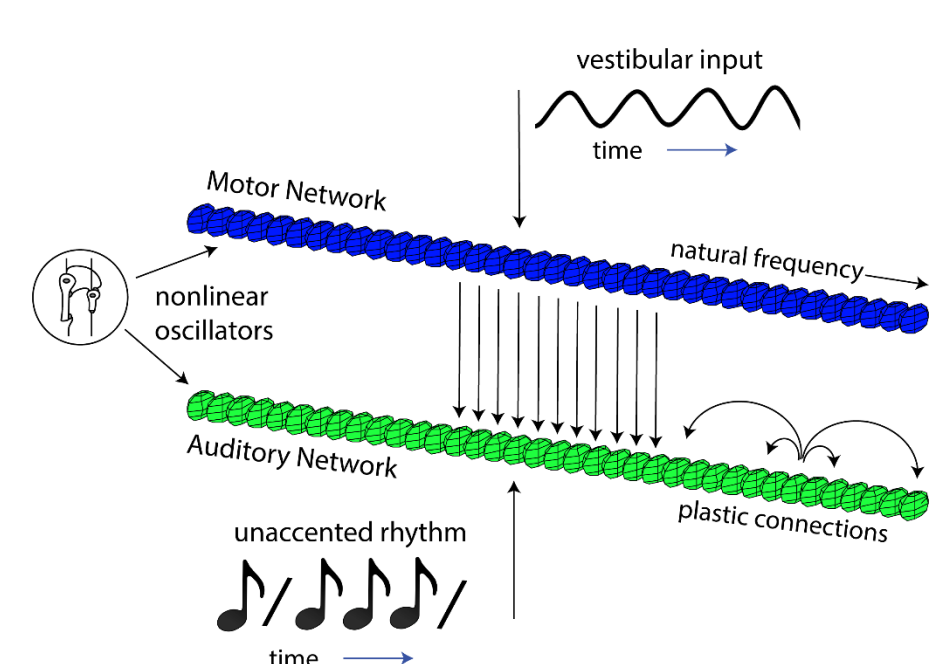
Background

Previous work suggests that auditory-vestibular interactions, which emerge during bodily movement to music, can influence the perception of musical structure, such as the perception of accented beats in musical rhythm¹⁻³. In a seminal study on the development of musical rhythm, Phillips-Silver & Trainor (2005) found that periodic, maternal bouncing of 7-month-old infants to an unaccented rhythm influenced infants' listening preferences for accented rhythms that matched the rate of maternal bouncing, suggesting that auditory-vestibular interactions shape rhythm perception in infancy. Expanding a recent theoretical model of infant rhythm perception⁴, in the current study, we propose a dynamical systems model of auditory-vestibular interactions thought to underlie infants' listening preferences for accented rhythms. The model, featuring two neural networks of non-linear oscillators to represent developmentally nascent auditory and motor systems, was used to simulate the effect of maternal bouncing (e.g., vestibular input) on infants' listening preferences for dupe- and triple-accented rhythms. First, we demonstrate that simultaneous auditory-vestibular training shaped the model's response to musical rhythm online, enhancing vestibular-related frequencies in the model's oscillatory activity. Next, we demonstrate that simultaneous auditory-vestibular training, relative to models that received auditory- or vestibular-only training, facilitated neural plasticity, producing stronger connections between network oscillators during a period of unsupervised learning. Finally, we show that models which received simultaneous auditory-vestibular training, but not models that received auditory-only or vestibular-only training, "preferred" rhythmic frequencies related to their "bouncing," resonating more strongly at frequencies related to the combined auditory-vestibular stimulation. This finding is qualitatively similar to infants' preferences for accented rhythms that matched the rate of maternal bouncing to an unaccented rhythm.

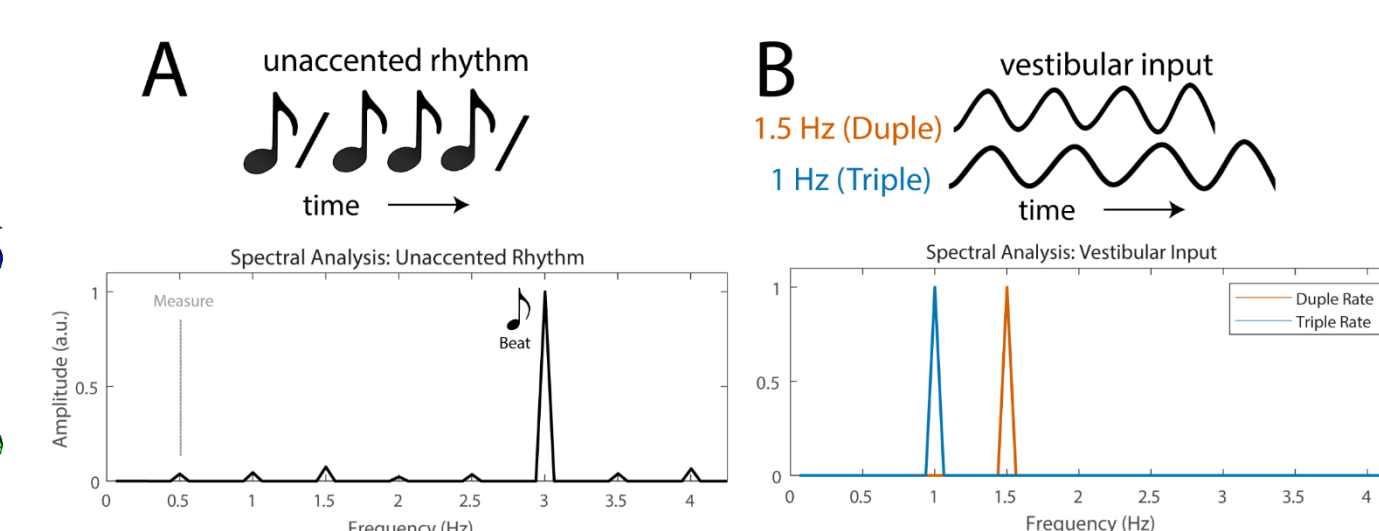


Methods

Model Architecture



Training Stimuli



• Notation and frequency-domain representations of the (A) Unaccented Rhythm (i.e., auditory input) (B) and Duple- and Triple-Rate Maternal Bouncing (i.e., vestibular input modeled as sinusoidal forcing).

• The model consisted of two multi-frequency oscillatory neural networks, representing developmentally nascent auditory and motor systems, with a Hebbian plasticity rule in the auditory network.

Training and Test Procedures

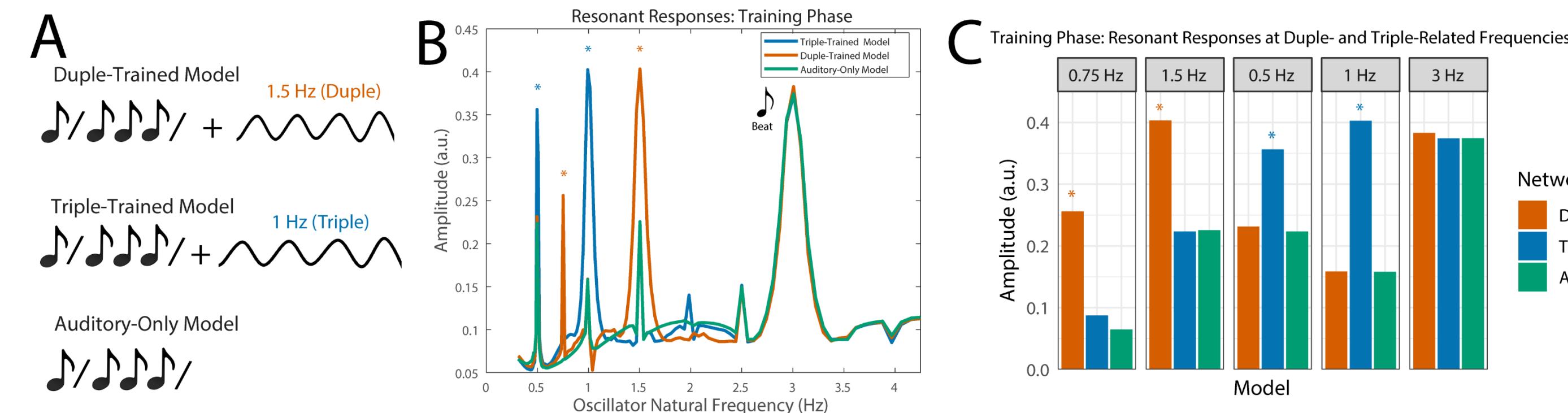
Network	Training Phase	Test Phase
Auditory Network		
Motor Network		

Training Regimes

Model: Training	Auditory Training	Vestibular Training
Duple-trained model	X	X
Triple-trained model	X	X
Auditory-only model	X	-
Duple-only model	-	X
Triple-only model	-	X

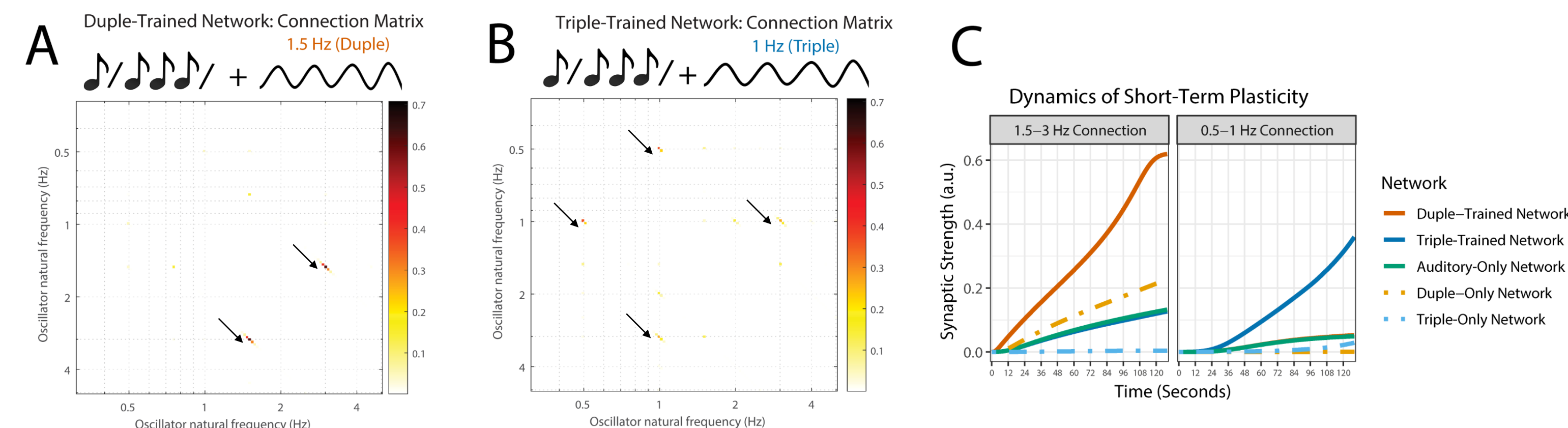
• To assess the effects of combined auditory-vestibular training on the model's behavior, two models were trained on simultaneous auditory-vestibular input (e.g., duple- and triple-trained models), similar to the infants in Phillips-Silver & Trainor (2005).
 • To assess the independent effects of auditory or vestibular training on the model's behavior, three control models were trained on auditory-only input (e.g., auditory-only model) or vestibular-only input (e.g., duple- and triple-only models).

Vestibular Input Shapes the Neural Response to Musical Rhythm



(A) Training stimuli for the duple-trained, triple-trained, and auditory-only models. (B) Resonant responses in the auditory network of the duple-trained (red), triple-trained (blue), and auditory-only (green) models during the final half of the training procedure. (C) Resonant-response amplitude values that were extracted from the resonant responses in the duple-trained (red), triple-trained (blue), and auditory-only (green) models for duple-related (0.75, 1.5 Hz, marked with red asterisks), triple-related (0.5, 1 Hz, marked with blue asterisks), and the beat frequency (3 Hz). Here, the duple-trained model produced larger responses at duple-related frequencies (0.75, 1.5 Hz) relative to the other models, while the triple-trained models produced larger responses at triple-related frequencies (0.5, 1 Hz) relative to the other models, suggesting that simultaneous auditory-vestibular training enhanced oscillatory responses to musical rhythm specifically at frequencies (i.e., harmonic and subharmonic frequencies) related to the rate of vestibular input.

Auditory-Vestibular Training Facilitates Neural Plasticity

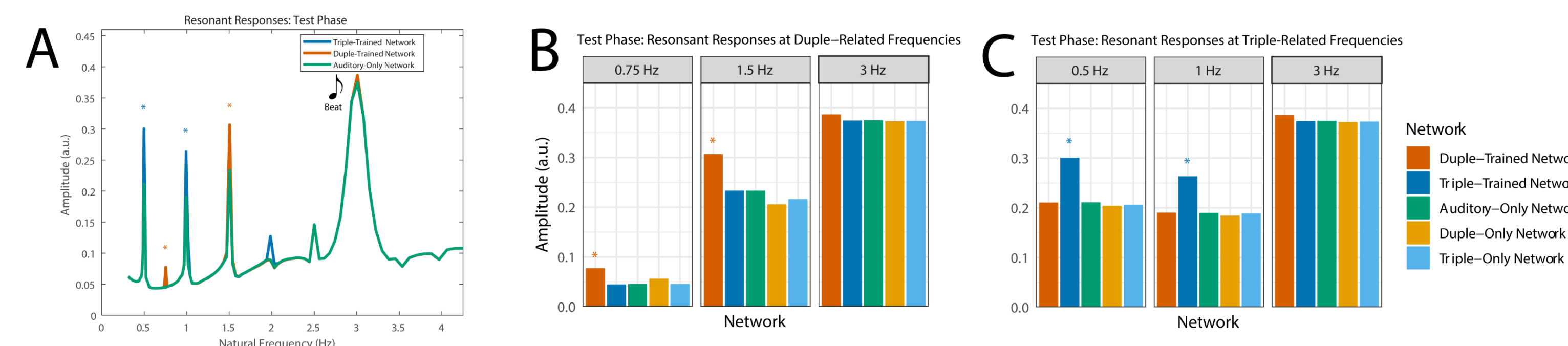


(A) Connection matrix of the duple-trained model. The duple-trained model, which was trained on an unaccented rhythm and duple-rate vestibular input, learned strong, bi-directional connections between 1.5 – 3 Hz, reflecting its rate of bouncing (1.5 Hz) and the beat level (3 Hz) of the rhythm.

(B) Connection matrix of the triple-trained model. The triple-trained model, a model trained on an unaccented rhythm and triple-rate vestibular input, exhibited more distributed learning, with bi-directional connections emerging between 0.5 – 1 Hz and 1 – 3 Hz, reflecting its rate of bouncing (1 Hz) and the beat and measure level (3 Hz, 0.5 Hz, respectively) of the rhythm.

(C) The time-course of learning oscillator connections at the strongest duple-related frequencies (1.5 – 3 Hz) and triple-related frequencies (0.5 – 1 Hz). (C, Left) The duple-trained network learned the strongest synaptic connection between network oscillators at 1.5 and 3 Hz, across all models. (C, Right) The triple-trained network learned the strongest synaptic connection between network oscillators at 0.5 and 1 Hz, relative to all models.

Auditory-Vestibular Models Preferred Vestibular-Related Rhythmic Frequencies



(A) Resonant responses (i.e., average oscillatory activity) for the duple-trained (red), triple-trained (blue), and auditory-only (green) models during the final half of the test procedure. Here, the duple- and triple-trained models resonated strongly at frequencies related to their vestibular training in response to the unaccented rhythm, suggesting that the models preferred rhythmic frequencies related to their bouncing.

(B) Resonant-response amplitude values for all models for duple-related frequencies during the test phase. Here, the duple-trained model produced the largest resonant responses at duple-related frequencies (0.75, 1.5 Hz; red asterisks) in response to the unaccented rhythm during the test phase, suggesting that the duple-trained model preferred rhythmic frequencies related to its bouncing.

(C) Resonant-response amplitude values for all models for triple-related frequencies during the test phase. Here, the triple-trained model produced the largest resonant responses at triple-related frequencies (0.5, 1 Hz; blue asterisks) in response to the unaccented rhythm during the test phase, suggesting that the triple-trained model preferred rhythmic frequencies related to its bouncing.

Summary and Interpretations

- Auditory-vestibular training shaped the neural response to musical rhythm, enhancing oscillatory activity at vestibular-related frequencies.
- Auditory-vestibular training facilitated neural plasticity, producing stronger connections between oscillators, relative to auditory- or vestibular-only training.
- Auditory-vestibular training engendered a preference in the model for vestibular-related rhythmic frequencies:
 - Similar to infants' preference for accented rhythms that matched the rate of maternal bouncing (Phillips-Silver & Trainor, 2005).

Future Directions

- Towards a general framework of rhythm development using Neural Resonance Theory^{4,5}.
 - Model the effects of short- and long-term training on rhythm perception-action⁶
 - Model tempo preferences across the lifespan⁷
 - Implement multiple timescales of learning into the model

References

1. Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the Beat: Movement Influences Infant Rhythm Perception. *Science (New York, N.Y.)*, 308(5727), 1430.
2. Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition*, 105(3), 533–546. <https://doi.org/10.1016/j.cognition.2006.11.006>
3. Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, 67(1), 94–102. <https://doi.org/10.1016/j.bandc.2007.11.007>
4. Tichko, P., & Large, E. W. (2019). Modeling infants' perceptual narrowing to musical rhythms: neural oscillation and Hebbian plasticity. *Annals of the New York Academy of Sciences*, 1453(1), 125–139.
5. Large, E., Herrera, J., & Velasco, M. (2015). Neural Networks for Beat Perception in Musical Rhythm. *Frontiers in Systems Neuroscience*, 9(November), 159.
6. Zhao, T. C., & Kuhl, P. K. (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proceedings of the National Academy of Sciences*, 201603984.
7. Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, 77(3), 251–288.

Model Equations

$$(1a) \tau_1 \dot{z}_{1t} = z_{1t} \left(\alpha + i2\pi + \beta_1 |z_{1t}|^2 + \frac{\beta_2 |z_{2t}|^4}{1 - |z_{1t}|^2} \right) + x(t) + \sum_{j \neq 1} c_{1j} z_{1t}^{k_{1j}} z_{2t}^{m_{1j}-1} + \sum_j d_{1j} \frac{z_{2t}}{1 - z_{2t}} \cdot \frac{1}{1 - \bar{z}_{2j}} \cdot \frac{1}{1 - \bar{z}_{1j}}$$

$$(1b) \tau_2 \dot{z}_{2t} = z_{2t} \left(\alpha + i2\pi + \beta_1 |z_{2t}|^2 + \frac{\beta_2 |z_{1t}|^4}{1 - |z_{2t}|^2} \right) + y(t)$$

Equations (1a, 1b) describe the dynamics of neural oscillators in the auditory (1a) and motor networks (1b). Here, z_{1t} and z_{2t} are complex-valued state variables whose real part represents the excitatory activity and whose imaginary part represents the inhibitory activity of the P^{th} neural oscillator in the auditory network, 1, and motor network, 2, respectively. The natural frequencies of the P^{th} oscillator in the auditory network, 1, and motor network, 2, is given by $f_{1t} = k_{1t}/\tau_{1t}$ and $f_{2t} = k_{2t}/\tau_{2t}$, respectively. When $\beta_1 < 0$, the endogenous activity of network oscillators is governed by the α parameter, where $\alpha = 0$ is the critical point. For $\alpha < 0$, the network oscillators will exhibit damped oscillation, while for $\alpha > 0$, the network oscillators will spontaneously oscillate via an Andronov-Hopf bifurcation. The β s reflect nonlinear damping parameters. Finally, c_{ij} represents input from the synaptic connection between the j th oscillator to the i th oscillator in the auditory network that are in a mode-locked relationship (i.e., k_{ij} integer-ratio relationships), while d_{ij} represents oscillatory input from efferent connections that connect the motor network to the auditory network. $x(t)$ and $y(t)$ are time-varying inputs to the auditory and motor network, which reflect a musical rhythm and maternal bouncing, respectively.

$$(2) \tau \dot{c}_{ij} = c_{ij} \left(\lambda_{ij} + \mu_{ij} |c_{ij}|^2 + \frac{\epsilon_c \mu_{ij} |c_{ij}|^4}{1 - \epsilon_c |c_{ij}|^2} \right) + \kappa \sqrt{\epsilon_c} k_{ij} + m_{ij} - 2 z_{ij}^{m_{ij}} z_{ij}^{k_{ij}}$$

Equation (2) describes the dynamics of Hebbian learning in the auditory network of the model. Here, c_{ij} is a complex variable that represents the amplitude and phase of the synaptic connection between the j th and i th oscillator in the auditory network. The parameters λ , μ , ϵ_c , κ , and ϵ_c determine the dynamics of the plasticity for oscillators that are near resonant frequency relationships, $m_{ij} = k_{ij}$. The unsubscripted 1 represents the global learning timescale. We set τ to a small value (i.e., faster timescale) to simulate short-term plasticity thought to arise from auditory-vestibular training.