

RESEARCH ARTICLE

Tactile stimulations reduce or promote the segregation of auditory streams: psychophysics and modeling

Farzaneh Darki ^{*}, James Rankin, Piotr Słowiński

Department of Mathematics and Statistics, Faculty of Environment, Science and Economy, University of Exeter, Exeter, United Kingdom

^{*} f.darki2@exeter.ac.uk

Abstract

Auditory stream segregation plays a crucial role in understanding the auditory scene. This study investigates the role of tactile stimulation in auditory stream segregation through psychophysics experiments and a computational model of audio-tactile interactions. We examine how tactile pulses, synchronized with one group of tones (high- or low-frequency tones) in a sequence of interleaved high- and low-frequency tones (ABA-triplets), influence the likelihood of perceiving integrated or segregated auditory streams. Our findings reveal that tactile pulses synchronized with a single tone sequence (either the A-tone or B-tone sequence) enhance perceptual segregation, while pulses synchronized with both tone sequences promote integration. Based on these findings, we developed a dynamical model that captures interactions between auditory and tactile neural circuits, including recurrent excitation, mutual inhibition, adaptation, and noise. The proposed model shows excellent agreement with the experiment. Model predictions are validated through psychophysics experiments. In the model, we assume that selective tactile stimulation dynamically modulates the tonotopic organization within the auditory cortex. This modulation facilitates segregation by reinforcing specific tonotopic responses through single-tone synchronization while smoothing neural activity patterns with dual-tone alignment to promote integration. The model offers a robust computational framework for exploring cross-modal effects on stream segregation and predicts neural behavior under varying tactile conditions. Our findings imply that cross-modal synchronization, with carefully timed tactile cues, could improve auditory perception with potential applications in auditory assistive technologies aimed at enhancing speech recognition in noisy settings.

OPEN ACCESS

Citation: Darki F, Rankin J, Słowiński P (2025) Tactile stimulations reduce or promote the segregation of auditory streams: psychophysics and modeling. *PLoS Comput Biol* 21(8): e1012701. <https://doi.org/10.1371/journal.pcbi.1012701>

Editor: Ulrik R. Beierholm, Durham University, UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND

Received: December 5, 2024

Accepted: July 23, 2025

Published: August 18, 2025

Copyright: © 2025 Darki et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The source code for the model and the datasets generated and analyzed during the current study are available in the GitHub repository [farzaneh-darki/Darki2024_tactile](https://github.com/farzaneh-darki/Darki2024_tactile): https://github.com/farzaneh-darki/Darki2024_tactile.

Funding: FD and JR acknowledge support from an Engineering and Physical Sciences Research Council (EPSRC) Standard Grant (Healthcare Technologies), (EP/W032422/1). <https://www.ukri.org/councils/epsrc/> The funders had no

Introduction

The difficulty in following a single voice or conversation in noisy situations is one of the main frustrations of people with hearing impairment, even when fitted with modern hearing aids [1]. This difficulty arises from the challenge of separating a single sound stream

role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

from many competing sounds, a process known as auditory stream segregation [2,3]. Auditory stream segregation has been widely studied with the auditory streaming paradigm, an idealized stimulus for which two sequences of pure tones can be segregated into different streams [4,5]. Stream segregation has been shown to be influenced by cross-modal interaction [6]. For example, synchronizing visual flashes with tones in an ABA- sequence has been shown to promote segregation, suggesting that temporal alignment across sensory modalities can bias perceptual organization [7,8]. Similarly, visual distractors in multitone masking paradigms have been found to enhance auditory selective attention and improve target detection within complex soundscapes [9]. Tactile cues have also been shown to improve the ability to segregate interleaved melodies, even under challenging listening conditions [10]. Tactile stimuli have the capacity to improve speech recognition in noise [11–13]. However, our understanding of the underlying neural dynamics of this process is still limited.

Sensory substitution theory provides a compelling framework for understanding how one sensory modality can compensate for a deficit in one or augment another during complex perceptual tasks [14,15]. Auditory perception, for example, can be influenced by stimuli from other sensory modalities, such as vision [16]. Although the impact of visual input on auditory perception has been extensively studied, the influence of somatosensory (tactile) stimuli remains less well understood [17–19]. While research suggests that tactile inputs can modulate auditory perception [20,21], it is unclear which attributes of tactile and auditory signals (e.g., frequency, intensity, rhythm, timing, and duration) play the most critical role in enhancing or suppressing the combined sensory experience.

In [22], the influence of tactile distractors on the ability to discriminate the frequency and intensity of auditory tones was studied through a series of psychophysical experiments. The authors demonstrated that auditory frequency perception was systematically biased by tactile distractors: distractors at frequencies lower than that of the auditory tones induced larger bias effects than distractors at higher frequencies. They also observed that tactile distractors biased the intensity of auditory perception. The magnitude of this effect scaled with distractor intensity, but did not vary with distractor frequency [22]. These findings highlight the complex interplay between tactile and auditory modalities, raising questions about the nature of cross-modal interactions between touch and sound, as well as the specific neural mechanisms underlying audio-tactile interactions.

Multisensory integration is traditionally thought to occur in higher association cortices, from which multisensory signals are relayed to (subcortical) areas involved in planning and executing actions [23,24]. According to this view, multisensory integration occurs only after unisensory information has been thoroughly processed along its specific sensory hierarchy. However, recent results challenge this notion and suggest that multisensory interactions can occur in early sensory areas. In particular, fMRI [25–29] and electrophysiological studies in monkeys [30–32] found multisensory activations in brain areas considered unisensory, supporting the idea that multisensory integration can occur early in sensory processing through feedforward mechanisms, independent of attention (preattentive bottom-up mechanisms [25]).

The audiotactile cross-modal interaction could occur due to both direct and indirect pathways connecting auditory and somatosensory processing areas [33]. Some neurons within the primary auditory cortex (A1) respond to tactile stimuli, directly reflecting tactile processing in auditory regions [30,34]. Tactile pulses can influence tonotopic responses in the auditory cortex [35]. Although A1 is traditionally associated with processing sound frequencies, studies show that it can also respond to multisensory inputs [21,36], such as tactile, particularly when those stimuli are temporally or spatially relevant to auditory processing. Tactile stimulation, especially rhythmic pulses, can influence the timing and phase of neural oscillations in

A1 [37,38], potentially enhancing the synchronization of responses. This is more likely if the tactile input is closely synchronized with auditory input, as it can alter the salience of specific frequencies [39,40].

Mathematical modeling has advanced our understanding of bistable phenomena including auditory streaming [41,42] and tactile rivalry [43]. A variety of models have been proposed to account for the segregation of two streams of sounds [41,44]. Rankin et al. [45–47] proposed a neuromechanistic model that captures the complex dynamics of perceptual alternations, including the dependence of perceptual durations on parameters such as frequency differences and presentation rate. The model reproduces characteristic features, such as the log-normal distribution of perceptual durations and the build-up effect. The model replicates the periodic, pulsatile responses of A1 and its dependence on stimulus features, which are pooled as inputs to a downstream competition stage. The competition between units arises from a combination of mutual inhibition, adaptation, and additive noise mechanisms, which are thought to contribute to perceptual bistability at cortical stages [48–50].

Although multisensory integration has received significant attention within the neuroscience community, the number of mechanistic mathematical models of multisensory integration remains limited. To bridge this gap, we propose a novel computational model for audio-tactile integration that can be generalized to other cross-sensory interactions. By integrating neuromechanistic modeling with psychophysics experiments, we elucidate how tactile sensations influence the perception of multiple sound sources and the underlying neural computations driving audio-tactile interactions. The proposed model extends the mathematical framework of auditory streaming. Specifically, we use the hypothesis that tactile inputs enhance excitation in the tonotopic response within the primary auditory cortex to extend the model of interactions between primary (A1) and non-primary auditory cortices presented in [45]. We experimentally validate the model. Our model aligns with preattentive bottom-up mechanisms proposed in [25]. Our work not only provides a robust platform for exploring audio-tactile interactions but also sets the stage for investigating other cross-sensory integrations, enhancing our fundamental understanding of multisensory processing.

Materials and methods

Ethics Statement

Approval was obtained from the ethics committee of the University of Exeter (eEMPS000058). The procedures used in this study adhere to the tenets of the Declaration of Helsinki. Written formal consent was obtained from all individual participants included in the study.

Psychophysics Experiments. Six volunteers (3 male, mean age 33.83 ± 6.74 SD) were recruited for Experiment 1 and twelve volunteers (7 male, mean age 35.5 ± 10.72 SD) were recruited for Experiment 2 from the University of Exeter. Each gave written informed consent and received minor monetary compensation for participating in a 1-hour session. Participants were naive to the purpose of the study and did not self-declare any neurological or sensory disorders. Procedures were in compliance with guidelines for research with human participants and approved by the University of Exeter Research Ethics Committee.

Here we used the auditory streaming paradigm [51] in which participants listen to a sequence of interleaved high- and low-frequency tones repeated in ABA- triplets (“A” and “B” tones, “-” silent gap) (Fig 1A). They were instructed to report the “integrated” percept when they perceived these sequences either as a single integrated stream (A B A - A B A - A B A -), and the “segregated” percept when they heard two distinct streams: one consisting of only A tones and the other of only B tones (concurrent: A - A - A - A - A - and - B - - - B - - - B - -) (Fig 1B). Participants sat in a sound-isolated booth and attended to auditory stimuli

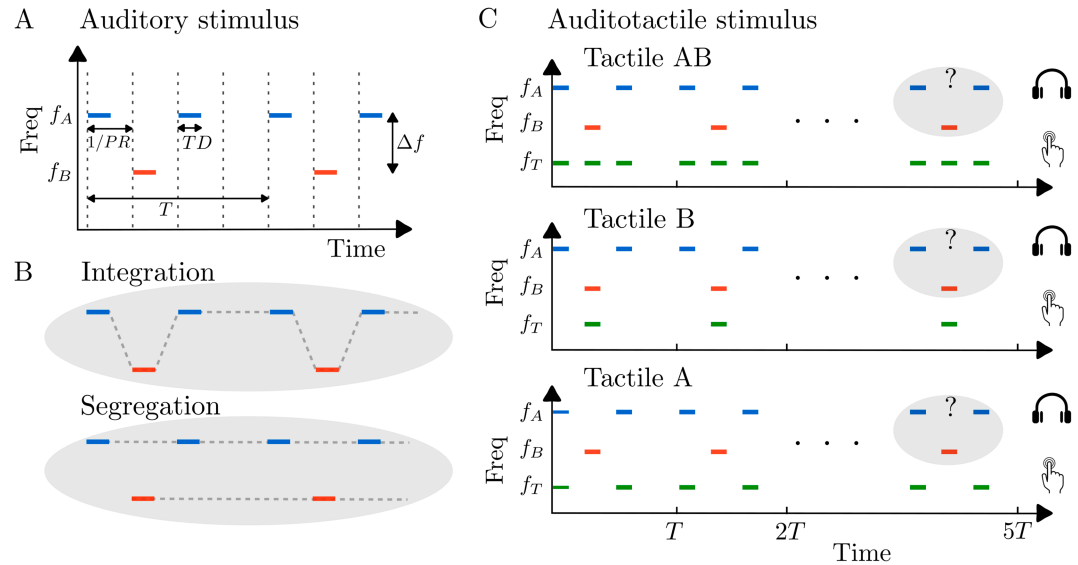


Fig 1. Auditory and tactile stimulus paradigm with two possible percepts. **A:** Repeating ABA- triplet sequences are composed of higher-frequency pure tones (A) interleaved with lower-frequency pure tones (B), each with a duration TD , and separated by a frequency difference Δf . The time interval between successive tone onsets (indicated by dashed vertical lines) corresponds to the inverse of the presentation rate ($1/PR$). The “-” in ABA- represents a silent interval of duration $1/PR$. In this study, tone duration is set to $TD = 1/PR$, ensuring that the offset of an A tone aligns precisely with the onset of the subsequent B tone. **B:** The stimulus is perceived in one of two ways: either as an integrated single stream (ABA-ABA-) or as two segregated streams (A-A-A-A- and -B-B-). Each experimental trial consisted of five consecutive ABA- triplets, and participants were asked to report their perception as integrated or segregated for the last triplet of the 5-triplet sequence. **C:** In the Tact AB condition, tactile stimulations at f_T occur during both A- and B-tone intervals. In contrast, in the Tact B and Tact A conditions, tactile stimulations occur only during B-tone or A-tone intervals, respectively.

<https://doi.org/10.1371/journal.pcbi.1012701.g001>

while vibrotactile stimulators were attached to their left index finger. To ensure that participants fully understood these interpretations, auditory and visual demonstrations were provided. Participants were instructed to report their perceptions passively, without trying to favor one organization over the other. They used keyboard presses to indicate their perceptual responses.

The auditory stimuli consisted of five consecutive ABA- triplets, where each triplet included 125 ms pure tones (A and B) followed by a 125 ms silence (“-”), resulting in a total triplet duration of 500 ms. Cosine-squared ramps of 10 ms were applied to the onset and offset of each tone to ensure smooth transitions and avoid acoustic artifacts. The higher frequency A tones were a variable Δf semitones (st) above the lower frequency B tones. A minimum 2 s interval between trials was used after which participants could run the next trial when ready.

Vibrotactile stimuli consisted of 125 ms sinusoidal vibratory pulses at 200 Hz. We used tactile stimulation synchronized with a subset of the tones in an ABA- triplet. The tactile pulse amplitude was calibrated prior to the experiment to be clearly perceptible but non-intrusive, based on pilot testing with naive participants. Three different tactile pulse timings were considered (Fig 1C): one trial with tactile pulses synchronized with A tones (Tact A), one with tactile pulses synchronized with B tones (Tact B), and one with tactile pulses synchronized with both A and B tones (Tact AB). Additionally, a trial without tactile stimulation (audio only, Tact Off) was included. The tactile pulse amplitude was fixed across all conditions to isolate the effect of timing and alignment. Participants were instructed to focus on auditory stimuli and report their perceptions based on the final auditory triplet presented in each trial.

The sequence of tones was played binaurally through Sennheiser HD 400 PRO headphones. We used miniature vibrotactile electromagnetic solenoid-type stimulators (18 mm diameter, Dancer Design factors [52]) driven by a tactile amplifier (Dancer Design Tac-tamp [52]) to deliver tactile stimuli. The voltages applied to the tactors for 200 Hz sinusoidal vibration was 3.38 V. In Experiment 1, each of the 6 participants completed 240 trials, consisting of 20 repetitions for each of three tactile conditions (Tact B, Tact AB, and Tact Off), combined with four levels of frequency difference ($\Delta f = \{3, 4, 5, 6\}$ st). We used a 6×6 Latin square design with 40 randomized and unique grids so that the order of conditions for each participant in a block of trials was counterbalanced within/across participants and block repetitions (one block of 6 participants and 40 repetitions). In Experiment 2, each of the 12 participants completed 160 trials, consisting of 40 repetitions for each of the four tactile conditions (Tact B, Tact AB, Tact Off, and Tact A) at a fixed frequency difference ($\Delta f = 4$ st). Here we used a 4×4 Latin square design with 120 randomized and unique grids (three blocks of 4 participants and 40 repetitions).

Statistical analysis. To investigate the effects of frequency difference (Δf) and tactile condition on the proportion of trials reported as segregated, we employed a generalized linear mixed model (GLMM) with a logit link function [53], accounting for repeated measures within participants. This statistical analysis method is useful for modeling non-normally distributed responses, such as binary outcomes (segregated, not segregated), and incorporates both fixed effects and a random intercept for participants to account for individual variability [54]. The fixed effects in the model estimate population-level coefficients, representing the relationship between predictor variables and the response, while random effects control for individual variations.

Different GLMMs were fitted to explore the main effects of Δf , tactile conditions, and their interaction. The GLMM provides coefficient estimates that indicate the strength and direction of the effect of predictor variables on the outcome; a positive coefficient indicates that an increase in the predictor variable raises the likelihood of the outcome occurring, while a negative coefficient suggests the opposite. The absolute value of the coefficient represents the strength of the relationship, with larger values indicating a stronger effect. The significance level of 0.05 is used throughout this paper. All statistical analyses were conducted in the statistical package *R*. To estimate Cohen's effect size [55], we first compute the odds ratio (OR) based on the observed frequencies of the binary outcome variable in Experiment 1. This odds ratio compares the odds of the event occurring in the test group to those in the control group. For the sample size estimation, we used *G*Power* [56] with a z-test for the difference between two independent proportions, employing an allocation ratio of 1 to achieve 80% power.

Mathematical model for audio-tactile interaction. To investigate the effect of tactile pulses on the segregation of auditory streams, we developed a mechanistic mathematical model that captures interactions within auditory and tactile neural circuits. The neuronal circuits for competition and perceptual encoding included in the model are assumed to be downstream and receive inputs from the primary auditory cortex (A1). Neuronal activity is represented by mean firing rates, with competitive interactions arising through excitatory and inhibitory connections, slow adaptation, synaptic depression, and intrinsic noise.

The model is described by the following system of first-order differential equations based on the model presented in Rankin et al. (2015) [45], which is considered a discrete idealization of a tonotopically organized array:

$$\begin{aligned}\tau_r \dot{r}_A &= -r_A + F(\beta_e d_A e_A - \beta_i r_B - 2\beta_i r_{AB} - g a_A + i_A + t_A + \chi_A), \\ \tau_r \dot{r}_{AB} &= -r_{AB} + F(\beta_e d_{AB} e_{AB} - \beta_i (r_A + r_B) - g a_{AB} + i_{AB} + t_{AB} + \chi_{AB}),\end{aligned}$$

$$\begin{aligned}
 \tau_r \dot{r}_B &= -r_B + F(\beta_e d_B e_B - \beta_i r_A - 2\beta_i r_{AB} - g a_B + i_B + t_B + \chi_B) \\
 \tau_a \dot{a}_A &= -a_A + r_A, \\
 \tau_a \dot{a}_{AB} &= -a_{AB} + r_{AB}, \\
 \tau_a \dot{a}_B &= -a_B + r_B, \\
 \tau_e \dot{e}_A &= -e_A + r_A, \\
 \tau_e \dot{e}_{AB} &= -e_{AB} + r_{AB}, \\
 \tau_e \dot{e}_B &= -e_B + r_B, \\
 \tau_d \dot{d}_A &= -d_A + (1 - \kappa r_A), \\
 \tau_d \dot{d}_{AB} &= -d_{AB} + (1 - \kappa r_{AB}), \\
 \tau_d \dot{d}_B &= -d_B + (1 - \kappa r_B).
 \end{aligned}
 \tag{1}$$

Here, r_k are the mean firing rates of each population, indexed by $k = \{A, AB, B\}$. Dynamics of each population also depends on, spike frequency adaptation variable a_k , the recurrent NMDA-excitation variable e_k , and the synaptic depression of excitatory connections d_k . τ_r , τ_a , τ_e and τ_d are synaptic time constants of respective variables. β_e is the strength of recurrent NMDA-excitation e_k , which is modulated by the slow synaptic depression d_k . β_i is the strength of instantaneous inhibition by the other populations r_k . Inhibition from the r_{AB} unit to the r_A and r_B units is twice as strong. g is the strength of spike-frequency adaptation, a_k . κ is the strength of the slow synaptic depression.

Each population shares a sigmoidal firing rate function F , given by

$$F(u) = \frac{1}{1 + \exp(k_F(-u + \theta_F))},
 \tag{2}$$

with threshold θ_F and slope k_F . The model is driven by periodic inputs that replicate the tonotopic responses in A1 to ABA- sequences [57]. i_k the tonotopic responses in the auditory cortex driving downstream neural populations are given by:

$$\begin{pmatrix} i_A \\ i_{AB} \\ i_B \end{pmatrix} = \Omega I = \begin{pmatrix} \omega(0) & \omega(\Delta f) \\ \omega(\Delta f/2) & \omega(\Delta f/2) \\ \omega(\Delta f) & \omega(0) \end{pmatrix} \begin{pmatrix} I_A \\ I_B \end{pmatrix}.
 \tag{3}$$

The spread of auditory I_A and I_B is defined by auditory weighting function $w(\Delta f) = M_a \exp\left(\frac{-\Delta f}{\sigma_a}\right)$, where σ_a is a spatial decay parameter and M_a is the pulse amplitude. The input terms I_k ($k = \{A, AB, B\}$) mimic the onset-plateau responses to pure tones in A1 with onset timescale α_1 , plateau timescale α_2 and peak to plateau ratio Λ_2 . Inputs are given by the following double α -function where $H(t)$ is the Heaviside function:

$$I_k(t) = H(t) \left[\frac{\exp(2)}{\alpha_1^2} t^2 \exp\left(\frac{-2t}{\alpha_1}\right) + \Lambda_2 \frac{\exp(2)}{\alpha_2^2} t^2 \exp\left(\frac{-2t}{\alpha_2}\right) \right].
 \tag{4}$$

The effect of tactile input on auditory tonotopic responses (through modulation of neural activity in A1) is considered additive with inputs t_k given by:

$$\begin{pmatrix} t_A \\ t_{AB} \\ t_B \end{pmatrix} = NT = \begin{pmatrix} \nu(0) & \nu(\Delta f) \\ \nu(\Delta f/2) & \nu(\Delta f/2) \\ \nu(\Delta f) & \nu(0) \end{pmatrix} \begin{pmatrix} T_A \\ T_B \end{pmatrix}.
 \tag{5}$$

Input terms T_A and T_B represent tactile pulses that are synchronized with tone A and tone B, respectively. These inputs are defined by the double α -function given in Eq 4. The spread of these tactile pulses across tonotopic locations is defined by the tactile weighting function $\nu(\Delta f) = M_{ik} \exp\left(\frac{-\Delta f}{\sigma_{ik}}\right)$, with parameters σ_{t1} and M_{t1} when tactile pulses are synchronized with one tone (Tact A or Tact B), and parameters σ_{t2} and M_{t2} when tactile pulses are synchronized with both A and B tones (Tact AB).

Additive noise is introduced with independent stochastic processes χ_A , χ_B and χ_{AB} and added to the inputs of each population. Input noise is modeled as an Ornstein-Uhlenbeck process:

$$\dot{\chi}_k = -\frac{\chi_k}{\tau_X} + \gamma \sqrt{\frac{2}{\tau_X}} \zeta_k(t), \quad (6)$$

where τ_X is the timescale, γ the strength and $\zeta(t)$ a white noise process with zero mean.

Values of all model parameters defined throughout this section are given in Table 1. Simulations were run in MATLAB using a standard Euler-Murayama time stepping scheme with a stepsize of 1 ms. Reducing this stepsize by a factor of 10 did not change the results.

Results

Effect of tactile stimulation on perceptual segregation across a range of frequency differences

In Experiment 1, we investigate the effect of tactile pulses on the ability to segregate sound streams. In the experiment, participants listen to a sequence of five auditory triplets, consisting of interleaved high- and low-frequency tones arranged in an ABA- pattern (Fig 1A). The frequency difference was varied across a range, including lower frequency differences where the proportion of perceptual segregation tends to be low. They report whether they perceive the tones as a single, integrated stream or as segregated into two streams, based on their perception of the final triplet. We used tactile stimulation synchronized with a subset of the tones in an ABA- triplet (Fig 1C). Data were collected from 6 participants, repeating 20 times each of 12 conditions (1440 observations in total), with each participant reporting their perception across various conditions and frequency differences. When tactile stimulation timing matches only the B tone sequence, the proportion of trials in which participants report perceiving the sounds as segregated increases. When the tactile pulse timing matches the A and B tones a bias towards integration emerges. This effect was observed over a range of frequency difference values (Fig 2A).

Table 1. Parameters of the model with their corresponding values used in the simulations.

Parameters	Values	Parameters	Values	Parameters	Values
M_a	0.95	β_e	0.85	τ_r	10 ms
σ_a	3.10 st	β_i	0.3	τ_a	1400 ms
M_{t1}	0.45	γ	0.075	τ_c	70 ms
σ_{t1}	1.2 st	κ	0.25	τ_d	3000 ms
M_{t2}	0.10	g	0.065	τ_x	100 ms
σ_{t2}	18 st	α_1	15 ms	k_F	12
Λ_2	1/6	α_2	82.5 ms	θ_F	0.2

<https://doi.org/10.1371/journal.pcbi.1012701.t001>

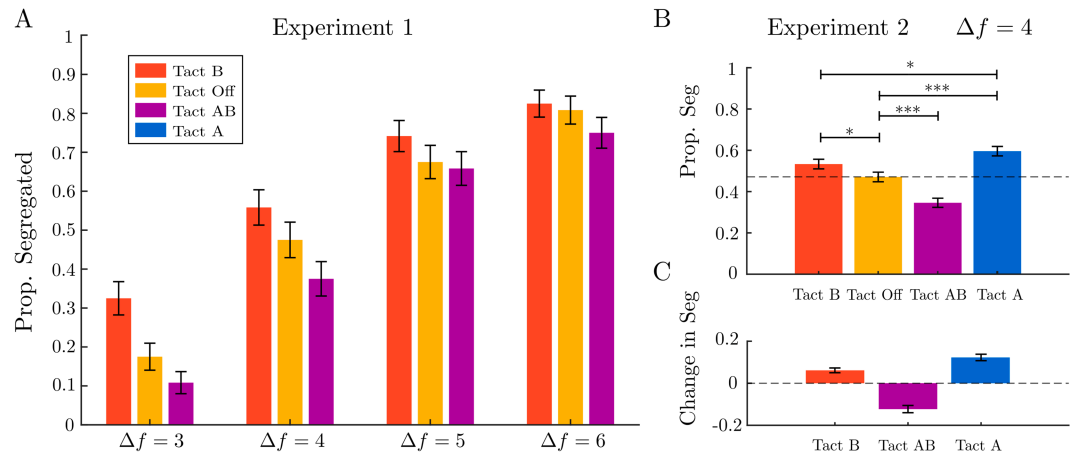


Fig 2. Effects of frequency difference, tactile conditions, and replicability. **A:** Experiment 1 (Combined effect of frequency difference and tactile conditions): Data were collected from six participants, each completing 20 repetitions at four frequency difference levels ($\Delta f = \{3, 4, 5, 6\}$) crossed with three tactile conditions, Tact B, Tact Off, Tact AB. Participants reported their perception across various tactile conditions and frequency differences. The proportion of trials in which the final (5th) triplet was reported as segregated increases with Δf and increases (decreases) in Tact B (Tact AB) relative to Tact Off. Error bars show the standard error of the proportion segregated. **B:** Experiment 2 (replicability, model validation): This experiment examines a fixed frequency difference ($\Delta f = 4$) and includes an additional tactile condition in which tactile pulses are aligned with tone A in ABA triplets. Compared to the reference condition Tact Off, the condition Tact B had a marginally significant positive effect, while Tact A and Tact AB exhibited significant positive and negative effects, respectively. Participants' perception of segregation was significantly lower in Tact B compared to Tact A. **C:** The change in proportion segregated compared to reference condition Tact Off.

<https://doi.org/10.1371/journal.pcbi.1012701.g002>

To investigate the relationship between frequency difference (Δf) and tactile condition and their effect on the perception of segregation in participants, we used a generalized linear mixed model (GLMM) analysis. The fixed effects of Δf and tactile condition and their interaction were examined to assess how they influenced the likelihood of segregation. A random intercept for participants was included to account for individual variability. The analysis revealed significant fixed effects for both Δf and tactile conditions. Specifically, the coefficient for Δf (Coeff. estimate = 1.01, $p < 0.001$) indicates a strong positive relationship between frequency difference and the likelihood of perceiving segregation, demonstrated in other studies without tactile stimulation [3,4,45]. Additionally, Tact B showed a positive effect (Coeff. estimate = 1.28, $p = 0.047$), while Tact AB exhibited a non-significant negative effect (Coeff. estimate = -0.64, $p = 0.35$). Notably, the interaction terms between Δf and the tactile conditions did not reach statistical significance, suggesting that tactile conditions have a consistent effect and do not vary significantly across different frequency differences (Table 2).

We also conducted statistical analyses on two reduced models that separately examined the Δf :Tact AB and Δf :Tact B interactions. In both cases, the interaction terms remained non-significant. Consequently, we adopted a simplified model excluding the interaction terms. This model revealed significant main effects for Δf (Coeff. estimate = 0.96, $p < 0.001$) and both tactile conditions Tact B (Coeff. estimate = 0.42, $p = 0.005$) and Tact AB (Coeff. estimate = -0.32, $p = 0.033$); see Table 3.

Analyzing the data at a specific frequency difference ($\Delta f = 4$) showed that Tact B had a positive effect on segregation perception compared to Tact Off (Coeff. estimate = 0.36, $p = 0.179$), whereas Tact AB exhibited a negative effect (Coeff. estimate = -0.45, $p = 0.103$). However, neither effect reached statistical significance (Table 4). The odds ratios of Tact AB and Tact B compared to the reference level of Tact Off are equal to 0.64 and 1.44, corresponding to

Table 2. The coefficients of GLMM of the effects of frequency difference (Δf) and tactile conditions on the segregation perception and their interactions while accounting for random effects of participants. Significant fixed effects are indicated with (*) or (*). The intercept and Δf are showing highly significant effects. The interaction between Δf and tactile conditions is not statistically significant. The model resulted in AIC = 1636.6 and BIC = 1673.5.**

Variable	Estimate	Std. Error	z value	P-value	
(Intercept)	-4.36	0.50	-8.72	<.001	***
Δf	1.01	0.10	9.70	<.001	***
Tact AB	-0.64	0.68	-0.94	0.350	
Tact B	1.28	0.64	1.99	0.047	*
Δf :Tact AB	0.07	0.15	0.47	0.638	
Δf :Tact B	-0.20	0.14	-1.39	0.166	

<https://doi.org/10.1371/journal.pcbi.1012701.t002>

Table 3. The coefficients of GLMM of the effects of frequency difference (Δf) and tactile conditions on the segregation perception while accounting for random effects of participants. The model indicates a significant positive effect of Δf and varying effects of the tactile conditions, with Tact AB showing a significant negative estimate and Tact B indicating a significant positive effect. The model resulted in AIC=1636.3 and BIC = 1662.6. Lower Std. Error and slightly lower BIC = 1662.6 value suggest a better model fit compared to the model including interactions (BIC = 1673.5).

Variable	Estimate	Std. Error	z value	P-value	
(Intercept)	-4.16	0.33	-12.66	<.001	***
Δf	0.96	0.06	16.05	<.001	***
Tact AB	-0.32	0.15	-2.13	0.033	*
Tact B	0.42	0.15	2.81	0.005	**

<https://doi.org/10.1371/journal.pcbi.1012701.t003>

Table 4. The coefficients of GLMM of the influence of tactile conditions on perceptual segregation at a fixed frequency difference ($\Delta f = 4$) in Experiment 1. Fixed effects Tact AB, and Tact B reveal no statistically significant differences in perceptual segregation between the conditions.

Variable	Estimate	Std. Error	z value	P-value
(Intercept)	-0.10	0.32	-0.31	0.757
Tact AB	-0.45	0.27	-1.63	0.103
Tact B	0.36	0.27	1.34	0.179

<https://doi.org/10.1371/journal.pcbi.1012701.t004>

Cohen's *d* (log odds ratio) 0.25 and 0.20, respectively. We used the smaller effect size of 0.2 of the tactile stimulation in Experiment 1 to estimate the sample size of a follow-up experiment (Experiment 2). To detect an effect size of 0.2 with 80% power using a z-test for the difference between two independent proportions, G*Power [56] estimates a required sample size of at least 444 (equivalent to 12 participants completing 40 repetitions for each condition).

Results from Experiment 1 reveal that tactile pulses have an effect on the ability to segregate auditory streams. Tactile pulses synchronized with tone B promote segregation, while those synchronized with both A and B tones promote integration. This suggests a complex interaction between tactile and auditory modalities, where the timing and context of tactile input are crucial. In Experiment 2, we investigate tactile stimulation at a fixed frequency difference ($\Delta f = 4$, close to equidominance point) to better understand the effects of temporal alignment of tactile and auditory stimuli. To this end, Experiment 2 also includes an additional tactile condition in which tactile pulses are aligned with A tones in ABA- triplets (in this way participants are exposed to twice as many tactile stimuli as in Tact B condition). The results of the experiment are shown in Fig 2B. The change in the proportion segregated for each tactile condition compared to the reference condition Tact Off is depicted in Fig 2C. Values greater than zero indicate a bias towards segregation, while values less than zero indicate a bias towards integration.

Compared to the reference condition Tact Off, Tact AB had a significant negative effect (Coeff. estimate = -0.54 , $p < 0.001$), while Tact A (Coeff. estimate = 0.53 , $p < 0.001$) and Tact B (Coeff. estimate = 0.26 , $p = 0.048$) exhibited significant positive effect; see also Table 5. Compared to Tact B, Tact A demonstrates a significantly higher proportion of segregation (Coeff. estimate = 0.27 , $p = 0.046$).

Mechanistic mathematical model of tactile-induced bias in auditory stream segregation

The model includes three neuronal populations. Two of them pool inputs from A1 regions centered on the frequencies A and B. The third population receives input from an intermediate tonotopic location, approximately $(A + B)/2$; see Fig 3A). The specific frequency tones exhibit full amplitude at their respective tonotopic locations, gradually decaying as they spread across other tonotopic locations (decay function $\omega(\Delta f)$). Similarly, the effect of tactile pulses is assumed to diminish across tonotopic locations, as illustrated by the changes in inputs corresponding to Tact B (represented by dashed curves in Fig 3B). When tactile pulses are synchronized with B-tones, they enhance the tonotopic response at location B and also induce an excitatory effect on other locations (AB and A), though to a lesser extent. We assumed a similar pattern of excitation for tactile pulses synchronized with A tones (decay function $\nu(\Delta f)$). However, when tactile pulses were synchronized with both tones, we considered a decaying effect, modeled with a different profile, to account for any residual impact from one tactile pulse to the next, as they occur at a shorter temporal distance. The top panel in Fig 3C shows a model simulation of the recurrent excitation variables for each population without the tactile effect (Tact Off). When the central AB unit is active (integrated), peripheral units are suppressed through mutual inhibition. Increasing adaptation for AB raises the probability of noise-induced switching, leading to the activation and dominance of units A or B (segregated), which in turn suppresses the integrated (AB) unit. The panels below depict predicted percepts under various tactile conditions, showing that Tact B increases the segregation proportion, while Tact AB decreases it compared to the Tact Off condition.

To investigate the mechanism of the effect of tactile pulses on the segregation of auditory streams we checked if the model presented in the section “Mathematical model for audio-tactile interaction” can be used to reproduce experimental observations. To this end, we used a genetic algorithm optimization implemented in Matlab function `ga` to find parameters M_a, M_{t1}, M_{t2} and $\sigma_a, \sigma_{t1}, \sigma_{t2}$ of the $\omega(\Delta f)$ and $\nu(\Delta f)$ functions; used in Eqs (3 and 4), respectively. The other model parameters were fixed throughout the optimization procedure. We defined the cost function as the mean squared error between the proportions of segregation observed in experiments and the proportions of segregation obtained in simulations. Details of the optimization procedure can be found in the computer code associated with the paper.

Table 5. The coefficients of GLMM of the influence of tactile conditions on perceptual segregation at a fixed frequency difference ($\Delta f = 4$) in Experiment 2.

Variable	Estimate	Std. Error	z value	P-value	
Tact AB - Tact Off	-0.54	0.14	-4.00	<.001	***
Tact B - Tact Off	0.26	0.13	1.98	0.048	*
Tact A - Tact Off	0.53	0.13	3.95	<.001	***
Tact A - Tact B	0.27	0.13	-2.00	0.046	*
Tact B - Tact AB	0.80	0.14	5.92	<.001	***
Tact A - Tact AB	1.07	0.14	7.81	<.001	***

<https://doi.org/10.1371/journal.pcbi.1012701.t005>

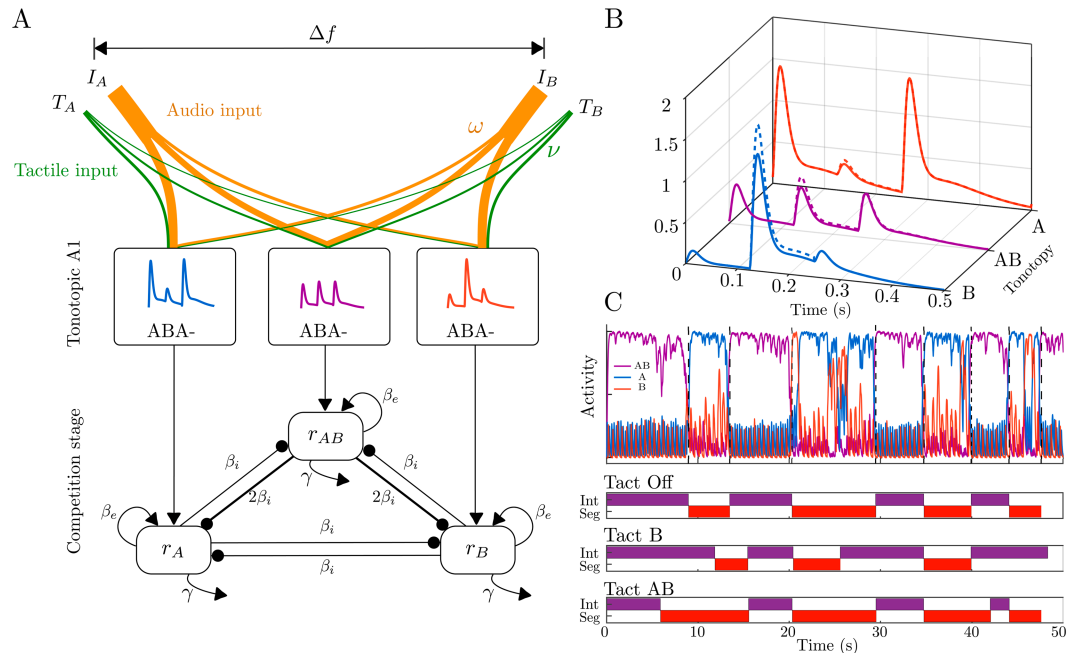


Fig 3. Model architecture, inputs, time course of model responses and predicted percepts. **A:** Neuromechanistic model with competition between units (r_A, r_{AB}, r_B) driven by inputs from three locations in tonotopic map in A1 (at A, B and a location in between). The competition stage is located downstream of (and takes input from) A1, with mutual inhibition between units (β_i), recurrent excitation (β_e), slow adaptation (γ) and noise-driven competition. **B:** Inputs to the respective populations r_A, r_{AB} and r_B for an ABA- triplet which shows the spread of inputs across the model's tonotopy. The A-tone and B-tone inputs (I_A and I_B) have full amplitude at their respective tonotopic locations and gradually decay as they spread across other tonotopic locations (solid curves). The effect of tactile pulses is also considered to decay across tonotopic locations (changes in inputs corresponding to Tact B represented by dashed curves). **C:** The top panel displays a model simulation of the recurrent excitation variables for each population without tactile effect (Tact Off). When the central AB unit is active (purple/ integrated), peripheral units are suppressed via mutual inhibition. Increasing adaptation in the AB unit raises the likelihood of noise-induced switching. The panels below show predicted percepts for different tactile conditions with tact B (Tact AB) increasing (decreasing) the segregation proportion compared to Tact Off condition.

<https://doi.org/10.1371/journal.pcbi.1012701.g003>

Results of the simulations for the optimal parameter values are shown in Fig 4. Experimental data are presented as bar plots, while results of the model simulations are shown as dashed curves with data points at different frequency differences ($\Delta f = \{3, 4, 5, 6\} st$) and varying tactile conditions (Fig 4A). The Δf -dependent profiles for the auditory input spread $\omega(\Delta f)$ (Fig 4C) and tactile input spread $\nu(\Delta f)$ (Fig 4D) were determined through the optimization algorithm. The decaying input function was estimated to have an amplitude of $M_a = 0.95 \pm 0.04$ and a slope of $\sigma_a = 3.10 \pm 0.12$. For the case where tactile pulses are aligned with one tone, the parameters are $M_{t1} = 0.45 \pm 0.12$ and $\sigma_{t1} = 1.2 \pm 0.22$. When tactile pulses are aligned with both tones, the amplitude is $M_{t2} = 0.10 \pm 0.04$ and the slope is $\sigma_{t2} = 18 \pm 7$. Confidence intervals are based on 20 repetitions of the fitting procedure. The effect of tactile input amplitude is smaller compared to auditory input, and the decay across the tonotopic representation is steeper in single-tone alignment than in auditory decay, while remaining almost uniform in dual-tone alignment.

We used model parameters estimated using data from Experiment 1 to run a new set of simulations that included tactile stimulus synchronized with tone A. We compared the results of the new simulation with data from Experiment 2. Without any additional fitting, the model made an accurate quantitative prediction that the level of segregation under the new tactile

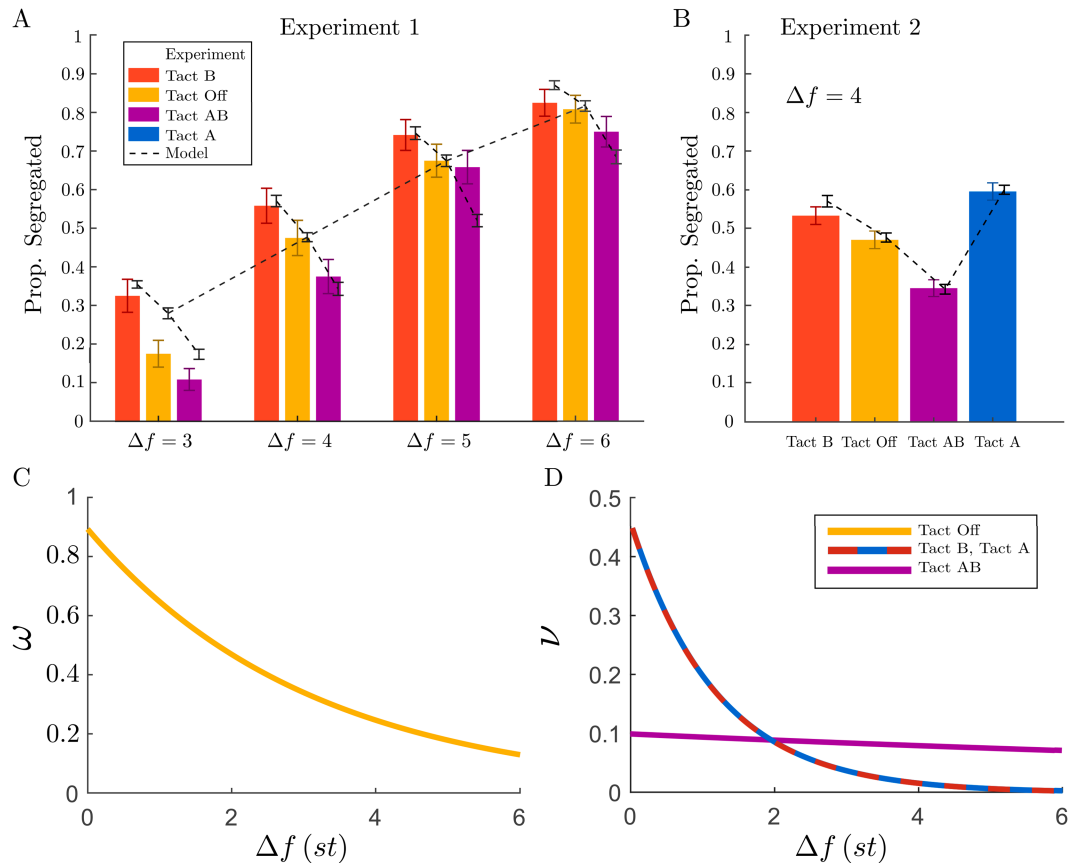


Fig 4. Comparison of experimental and simulated data across frequency differences and tactile conditions, with audio and tactile input spread profiles. A–B: Experimental data are presented as bar plots, while simulated data are shown as dashed curves with data points at different frequency differences ($\Delta f = 3, 4, 5, 6$ st) and under different tactile conditions. Error bars show the standard error of the proportion segregated. C–D: The Δf -dependent profiles for the auditory input spread $\omega(\Delta f)$ and tactile input spread $\nu(\Delta f)$ (modeled as exponential decays) were derived using an optimization algorithm minimizing the mean squared error between experimental and computational data from Experiment 1. These weighting functions were then used to compare computational model predictions with experimental data in Experiment 2.

<https://doi.org/10.1371/journal.pcbi.1012701.g004>

condition, Tact A, is higher than Tact B. Fig 4B illustrates excellent alignment between the model and the experimental data.

Discussion

Tactile pulses interfere with tonotopic responses

By synchronizing tactile pulses with one group of tones in the ABA- triplet sequence, we explored the perceptual shifts between integration (one stream) and segregation (two streams) based on the timing and pattern of tactile stimulation. Our findings imply that cross-modal synchronization can dynamically shift perceptual boundaries in auditory streaming. When tactile pulses coincided only with tone B, experiments demonstrated an increased likelihood of segregating the tones. When both A and B tones were paired with tactile pulses, participants perceived the sounds as a single stream more often. This suggests that tactile stimulation synchronized with a single group of tones in the sequence enhances perceptual distinction between tones, reinforcing the auditory separation needed for segregation, while alignment of

both tone sequences and tactile pulses provides a unified cross-modal cue, leading the auditory system to interpret the sequence as more cohesive. Tactile conditions have a consistent effect across various frequency differences suggesting that tactile timing relative to sound structure is a strong factor in stream segregation.

Tonotopic modulation and temporal averaging

In our model, tactile pulses modulate auditory perception by altering tonotopic responses in the auditory cortex. In the primary auditory cortex, neurons that respond to similar frequencies are organized in clusters along the tonotopic map [58]. Excitatory neurons activate neighboring frequency-tuned neurons, while inhibitory interneurons create lateral inhibition, sharpening frequency specificity and limiting spread across the map [59]. When tactile pulses align with a single group of tones, they appear to selectively boost responses in regions associated with that tone, with the effect decaying across the tonotopic map. This supports segregation by enhancing contrast in tonotopic representation. When both tones are paired with tactile stimuli, broader, more uniform activation across tonotopic areas reduces this contrast, promoting integration.

The increased segregation observed when tactile pulses align with tones A in the ABA-triplets, compared to B tones, is particularly notable due to the presence of two tactile pulses instead of one. With two tactile stimuli per ABA triplet, the tactile impact curve decays as sharply as before. However, because the tactile stimuli are now repeated twice within a triplet, they reinforce the auditory processing of tone A, leading to a more pronounced segregation effect.

We hypothesize that the observed effect could be explained through neural mechanisms in the auditory cortex. When tactile pulses synchronize with auditory tones, they interact with tonotopic circuits, activating frequency-tuned neurons and spreading in a balanced manner controlled by excitation and inhibition. When tactile pulses activate neurons associated with both tones, the individual effects integrated in time lead to a combined activation in the neural responses that do not decay as sharply across the tonotopic map. This could be due to recurrent excitation, where the remaining activity from one pulse keeps the neural population primed, leading to a more sustained response. The combination of effects at both tonotopic locations could create a stable perceptual outcome that enhances integration, as the stimulation is spread out more evenly [60].

Alternatively, the experimental observations could be explained by temporal averaging of excitatory and inhibitory effects. Such averaging would lead to the tactile impact curve that is smaller and smoother when tactile pulses align with both tones versus just one. When tactile pulses are synchronized with both tones, the excitatory and inhibitory responses may partially overlap and combine across time, leading to a more balanced, averaged response across the tonotopic map. This smoothing effect would reduce the peak strength of activation but maintain it across a wider spatial area, resulting in a smaller yet more stable effect. In contrast, with tactile pulses aligned with one tone, the activation is enhanced for that specific frequency region. Such localized amplification could explain the stronger but spatially limited effect in the one-tone condition.

Additive tactile effect and pattern of tactile pulses

Using the statistical GLM model we observed no significant interaction between frequency difference (Δf) and tactile pulses. However, an interaction effect with a small effect size may still exist (particularly in the case of Tact B) but could remain undetected given the statistical power of our study. This shows that within the limits of our study's power, the model

does not require the inclusion of multiplicative terms. Additionally, the Δf -dependent additive term results in a small interaction effect. This approach provides a simplified framework to capture the subtle tactile effects without introducing complex interaction terms. For the current dataset, an additive model appears sufficient. However, we cannot entirely rule out the presence of a subtle interaction that might be detected with a larger sample. This could be investigated further in future studies.

Our model parameter fitting approach allows us to estimate the strength of Δf -dependent additive effects and assess the degree of Δf dependency across different tactile conditions. The results from parameter fitting through optimization indicate that these effects are stronger and more Δf -dependent in the case of Tact B, compared to the Tact AB effect, which appears weaker and less dependent on Δf (see Fig 4). These findings are consistent with the visual interaction effect of Tact B observed in Fig 2.

Neural substrates for cross-modal influences on auditory streaming

Earlier studies using MEG [61,62] and intracranial EEG [63] have highlighted the potential role of mutual inhibition and selective attention in auditory streaming, suggesting that attention may modulate streaming through the sharpening of tonotopic representations. Our findings align with this framework, indicating that tactile pulses can serve as attention-enhancing cues that selectively reinforce tonotopic responses (top-down attentional mechanism). While our model emphasizes a bottom-up, feedforward modulation of tonotopic activity by tactile input, we acknowledge that attention-based enhancement of the attended tone stream via tactile pairing may also contribute to the observed perceptual biases. Future studies employing explicit manipulations of attention or incorporating distractor tasks could help disentangle the relative contributions of bottom-up and top-down mechanisms.

While our model is not biophysical and does not distinguish between cortical layers or neuronal types, it is still useful to consider potential biological substrates. Research shows that both intra-modal and cross-modal interactions often target superficial cortical layers, including the apical dendrites of pyramidal neurons in sensory cortex [64,65]. Within this framework, tactile input may influence auditory processing by increasing excitatory input to apical dendrites in frequency-specific columns of auditory cortex. This could modulate neural gain or bias perception toward tones that are synchronized with tactile input, aligning with our psychophysical findings and model behavior. Although our approach simplifies cortical architecture, the proposed model aligns with evidence that apical dendritic excitation integrates contextual and multisensory inputs. Future work could incorporate anatomically based, layered circuits to explore how dendritic processing shapes cross-modal auditory perception.

Model and experimental differences relative to prior work

Rankin et al. (2015) [45] proposed a neuromechanistic model capable of capturing the build-up effect through a gradual decrease in A1 response amplitudes applied to the first three triplets, thereby replicating the favor of integration during the build-up phase. Here we used a simplified version of this model [46] without amplitude modulation. This decision was driven by the current uncertainty surrounding how tactile inputs are processed during the build-up phase, and whether their influence strengthens or diminishes over time. Further experiments specifically designed to capture the influence of tactile pulses during the early triplets, and to assess whether this effect gradually changes across subsequent triplets, would help quantify the interaction of these effects.

One notable point of comparison between our findings and previous work arises when examining the proportion of segregation reported at intermediate frequency differences. In

Fig 2 of our study, participants reported segregation approximately 65% of the time at $\Delta f = 5$. This contrasts with the model presented in Rankin et al. (2015) [45], where equidominance between integration and segregation was observed near $\Delta f = 5$. This difference likely reflects both variations in experimental design (Rankin et al. [45] modeled long trials with full time-based dynamics, whereas we used short sequences with perceptual reports based on the final triplet) and differences in participant characteristics (e.g., age range, musical training, attentional strategies).

Model predictions and future work

Based on the current stage of our model, which is built upon the present dataset, we predict that tactile pulses aligned with either the first or last A tones would bias perception toward segregation. Our ongoing experiments investigate the effects of tactile pulses that are asynchronous with the auditory tones (specifically, when tactile pulses lead or lag the tones), and we are expanding the model to incorporate neural mechanisms that may underlie these effects. Tactile pulses delivered in the inter-tone gaps could either reduce interference or enhance temporal binding, depending on their proximity to the A and B tones and the number of repetitions in each triplet.

Fig 2 suggests that tactile stimulation may be more effective at promoting segregation at lower Δf values, and more effective at promoting integration at higher Δf values. This pattern might become even more pronounced at more extreme frequency differences beyond the range tested in our study. Future studies could investigate whether tactile cues exert their strongest influence near perceptual thresholds. Such investigations will help constrain future model development and clarify how temporal alignment and frequency separation jointly shape cross-modal perceptual organization.

Audiotactile interactions and implications for multisensory integration

Previous studies have demonstrated that tactile distractors can bias auditory frequency perception, shifting perceived frequencies toward those of the tactile pulses [22]. This bias is more pronounced with lower-frequency tactile distractors compared to higher frequencies. In this study, we specifically used tactile pulses with frequencies lower than the auditory tones. Future research could investigate the impact of higher-frequency tactile pulses on auditory perception to determine if they produce a different bias or interaction pattern, deepening our understanding of multisensory integration dynamics across varying frequency contexts.

Our experimental and mathematical modeling results reveal a spatio-temporal interaction where tactile input dynamically influences auditory organization through timing and tonotopic alignment. Specifically, our findings show that dual-tone alignment creates a broad, stable neural response, while single-tone alignment produces a sharper, more localized effect. This interaction suggests that tactile cues, precisely timed and placed, can be employed to fine-tune auditory processing. Such intricate control offers valuable insights into multisensory integration, especially for the design of devices for individuals with hearing impairments or difficulties with speech recognition in noisy environments. By strategically incorporating tactile feedback, our approach can improve and optimize the design of these devices, enhancing auditory clarity and speech recognition.

Conclusion

This study highlights the significant role of tactile pulses in modulating auditory perception, demonstrating that synchronized tactile input can influence the integration and segregation

of auditory streams through spatiotemporal interactions. These effects arise from interactions within tonotopic circuits, where excitatory and inhibitory dynamics shape the perceptual outcome. Single-tone alignment creates localized enhancement and dual-tone alignment generates broader, averaged responses. The consistency of tactile effects across varying frequency differences highlights the robustness of tactile-auditory interactions. These insights open pathways for future research to explore how tactile inputs of different frequencies influence auditory perception, further advancing our understanding of multisensory integration. Importantly, the ability to fine-tune auditory processing using precisely timed tactile cues has promising applications, particularly in assistive technologies. This work lays the foundation for innovative solutions that capitalize on multisensory integration to optimize auditory experiences.

Acknowledgments

We would like to thank Marc Goodfellow and Mark Fletcher for their valuable discussions and insightful suggestions during the early stages of this project.

Author contributions

Conceptualization: Farzaneh Darki, James Rankin, Piotr Słowiński.

Data curation: Farzaneh Darki.

Formal analysis: Farzaneh Darki.

Funding acquisition: James Rankin.

Investigation: Farzaneh Darki, James Rankin, Piotr Słowiński.

Methodology: Farzaneh Darki, James Rankin, Piotr Słowiński.

Project administration: James Rankin, Piotr Słowiński.

Resources: Farzaneh Darki, James Rankin.

Supervision: James Rankin, Piotr Słowiński.

Validation: Farzaneh Darki, Piotr Słowiński.

Visualization: Farzaneh Darki.

Writing – original draft: Farzaneh Darki, Piotr Słowiński.

Writing – review & editing: Farzaneh Darki, Piotr Słowiński.

References

1. Chung K. Challenges and recent developments in hearing aids. Part I. Speech understanding in noise, microphone technologies and noise reduction algorithms. *Trends Amplif.* 2004;8(3):83–124. <https://doi.org/10.1177/108471380400800302> PMID: 15678225
2. Snyder JS, Alain C. Toward a neurophysiological theory of auditory stream segregation. *Psychol Bull.* 2007;133(5):780–99. <https://doi.org/10.1037/0033-2909.133.5.780> PMID: 17723030
3. Moore BCJ, Gockel HE. Properties of auditory stream formation. *Philos Trans R Soc Lond B Biol Sci.* 2012;367(1591):919–31. <https://doi.org/10.1098/rstb.2011.0355> PMID: 22371614
4. van Noorden LPAS. Temporal coherence in the perception of tone sequences; 1975.
5. Fishman YI, Reser DH, Arezzo JC, Steinschneider M. Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hear Res.* 2001;151(1–2):167–87. [https://doi.org/10.1016/s0378-5955\(00\)00224-0](https://doi.org/10.1016/s0378-5955(00)00224-0) PMID: 11124464

6. Carlyon RP, Plack CJ, Fantini DA, Cusack R. Cross-modal and non-sensory influences on auditory streaming. *Perception*. 2003;32(11):1393–402. <https://doi.org/10.1068/p5035> PMID: 14959799
7. Marozeau J, Innes-Brown H, Grayden DB, Burkitt AN, Blamey PJ. The effect of visual cues on auditory stream segregation in musicians and non-musicians. *PLoS One*. 2010;5(6):e11297. <https://doi.org/10.1371/journal.pone.0011297> PMID: 20585606
8. Rahne T, Deike S, Selezneva E, Brosch M, König R, Scheich H, et al. A multilevel and cross-modal approach towards neuronal mechanisms of auditory streaming. *Brain Res*. 2008;1220:118–31. <https://doi.org/10.1016/j.brainres.2007.08.011> PMID: 17765207
9. Hausfeld L, Gutschalk A, Formisano E, Riecke L. Effects of cross-modal asynchrony on informational masking in human cortex. *J Cogn Neurosci*. 2017;29(6):980–90. https://doi.org/10.1162/jocn_a_01097 PMID: 28129050
10. Slater KD, Marozeau J. The effect of tactile cues on auditory stream segregation ability of musicians and nonmusicians. *Psychomusicol: Music Mind Brain*. 2016;26(2):162–6. <https://doi.org/10.1037/pmu0000143>
11. Schulte A, Marozeau J, Ruhe A, Büchner A, Kral A, Innes-Brown H. Improved speech intelligibility in the presence of congruent vibrotactile speech input. *Sci Rep*. 2023;13(1):22657. <https://doi.org/10.1038/s41598-023-48893-w> PMID: 38114599
12. Fletcher MD, Verschuur CA, Perry SW. Improving speech perception for hearing-impaired listeners using audio-to-tactile sensory substitution with multiple frequency channels. *Sci Rep*. 2023;13(1):13336. <https://doi.org/10.1038/s41598-023-40509-7> PMID: 37587166
13. Fletcher MD, Akis E, Verschuur CA, Perry SW. Improved tactile speech perception and noise robustness using audio-to-tactile sensory substitution with amplitude envelope expansion. *Sci Rep*. 2024;14(1):15029. <https://doi.org/10.1038/s41598-024-65510-6> PMID: 38951556
14. Auvray M, Myin E. Perception with compensatory devices: From sensory substitution to sensorimotor extension. *Cogn Sci*. 2009;33(6):1036–58. <https://doi.org/10.1111/j.1551-6709.2009.01040.x> PMID: 21585495
15. Auvray M, Harris LR. The state of the art of sensory substitution. *Multisens Res*. 2014;27(5–6):265–9. <https://doi.org/10.1163/22134808-00002464> PMID: 25693296
16. Schwartz J-L, Grimault N, Hupé J-M, Moore BCJ, Pressnitzer D. Multistability in perception: Binding sensory modalities, an overview. *Philos Trans R Soc Lond B Biol Sci*. 2012;367(1591):896–905. <https://doi.org/10.1098/rstb.2011.0254> PMID: 22371612
17. MacDonald J, McGurk H. Visual influences on speech perception processes. *Percept Psychophys*. 1978;24(3):253–7. <https://doi.org/10.3758/bf03206096> PMID: 704285
18. King AJ. Visual influences on auditory spatial learning. *Philos Trans R Soc Lond B Biol Sci*. 2009;364(1515):331–9. <https://doi.org/10.1098/rstb.2008.0230> PMID: 18986967
19. Opoku-Baah C, Schoenhaut AM, Vassall SG, Tovar DA, Ramachandran R, Wallace MT. Visual influences on auditory behavioral, neural, and perceptual processes: A review. *J Assoc Res Otolaryngol*. 2021;22(4):365–86. <https://doi.org/10.1007/s10162-021-00789-0> PMID: 34014416
20. Bolognini N, Papagno C, Moroni D, Maravita A. Tactile temporal processing in the auditory cortex. *J Cogn Neurosci*. 2010;22(6):1201–11.
21. Hoefer M, Tyll S, Kanowski M, Brosch M, Schoenfeld MA, Heinze H-J, et al. Tactile stimulation and hemispheric asymmetries modulate auditory perception and neural responses in primary auditory cortex. *Neuroimage*. 2013;79:371–82. <https://doi.org/10.1016/j.neuroimage.2013.04.119> PMID: 23664954
22. Yau JM, Weber AI, Bensmaia SJ. Separate mechanisms for audio-tactile pitch and loudness interactions. *Front Psychol*. 2010;1:160. <https://doi.org/10.3389/fpsyg.2010.00160> PMID: 21887147
23. Cappe C, Rouiller EM, Barone P. Multisensory anatomical pathways. *Hear Res*. 2009;258(1–2):28–36. <https://doi.org/10.1016/j.heares.2009.04.017> PMID: 19410641
24. Klemen J, Chambers CD. Current perspectives and methods in studying neural mechanisms of multisensory interactions. *Neurosci Biobehav Rev*. 2012;36(1):111–33. <https://doi.org/10.1016/j.neubiorev.2011.04.015> PMID: 21569794
25. Kayser C, Petkov CI, Augath M, Logothetis NK. Integration of touch and sound in auditory cortex. *Neuron*. 2005;48(2):373–84. <https://doi.org/10.1016/j.neuron.2005.09.018> PMID: 16242415
26. Foxe JJ, Schroeder CE. The case for feedforward multisensory convergence during early cortical processing. *Neuroreport*. 2005;16(5):419–23. <https://doi.org/10.1097/00001756-200504040-00001> PMID: 15770144
27. Macaluso E, Driver J. Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends Neurosci*. 2005;28(5):264–71. <https://doi.org/10.1016/j.tins.2005.03.008> PMID: 15866201

28. Schroeder CE, Foxe J. Multisensory contributions to low-level, 'unisensory' processing. *Curr Opin Neurobiol.* 2005;15(4):454–8.
29. Raji T, Ahveninen J, Lin F-H, Witzel T, Jääskeläinen IP, Letham B, et al. Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *Eur J Neurosci.* 2010;31(10):1772–82. <https://doi.org/10.1111/j.1460-9568.2010.07213.x> PMID: 20584181
30. Fu KMG, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, et al. Auditory cortical neurons respond to somatosensory stimulation. *J Neurosci.* 2003;23(20):7510–5.
31. Schroeder CE, Lindsley RW, Specht C, Marcovici A, Smiley JF, Javitt DC. Somatosensory input to auditory association cortex in the macaque monkey. *J Neurophysiol.* 2001;85(3):1322–7. <https://doi.org/10.1152/jn.2001.85.3.1322> PMID: 11248001
32. Schroeder CE, Smiley J, Fu KG, McGinnis T, O'Connell MN, Hackett TA. Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing. *Int J Psychophysiol.* 2003;50(1–2):5–17. [https://doi.org/10.1016/s0167-8760\(03\)00120-x](https://doi.org/10.1016/s0167-8760(03)00120-x) PMID: 14511832
33. Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, et al. Auditory-somatosensory multisensory processing in auditory association cortex: An fMRI study. *J Neurophysiol.* 2002;88(1):540–3. <https://doi.org/10.1152/jn.2002.88.1.540> PMID: 12091578
34. Schürmann M, Caetano G, Hlushchuk Y, Jousmäki V, Hari R. Touch activates human auditory cortex. *Neuroimage.* 2006;30(4):1325–31. <https://doi.org/10.1016/j.neuroimage.2005.11.020> PMID: 16488157
35. Yau JM, Olenczak JB, Dammann JF, Bensmaia SJ. Temporal frequency channels are linked across audition and touch. *Curr Biol.* 2009;19(7):561–6. <https://doi.org/10.1016/j.cub.2009.02.013> PMID: 19268591
36. Henschke JU, Noesselt T, Scheich H, Budinger E. Possible anatomical pathways for short-latency multisensory integration processes in primary sensory cortices. *Brain Struct Funct.* 2015;220(2):955–77. <https://doi.org/10.1007/s00429-013-0694-4> PMID: 24384580
37. Fu X, Riecke L. Effects of continuous tactile stimulation on auditory-evoked cortical responses depend on the audio-tactile phase. *Neuroimage.* 2023;274:120140. <https://doi.org/10.1016/j.neuroimage.2023.120140> PMID: 37120042
38. Ranjbar P, Wilson EC, Reed CM, Braida LD. Auditory-tactile integration: Effects of phase of sinusoidal stimulation at 50 and 250 Hz. *Int J Eng Technol Sci Innov.* 2016;1(2):209–29. PMID: 28626839
39. Lunghi C, Morrone MC, Alais D. Auditory and tactile signals combine to influence vision during binocular rivalry. *J Neurosci.* 2014;34(3):784–92. <https://doi.org/10.1523/JNEUROSCI.2732-13.2014> PMID: 24431437
40. Jagt M, Ganis F, Serafin S. Enhanced neural phase locking through audio-tactile stimulation. *Front Neurosci.* 2024;18:1425398. <https://doi.org/10.3389/fnins.2024.1425398> PMID: 39416951
41. Szabó BT, Denham SL, Winkler I. Computational models of auditory scene analysis: A review. *Front Neurosci.* 2016;10:524. <https://doi.org/10.3389/fnins.2016.00524> PMID: 27895552
42. Rankin J, Rinzel J. Computational models of auditory perception from feature extraction to stream segregation and behavior. *Curr Opin Neurobiol.* 2019;58:46–53. <https://doi.org/10.1016/j.conb.2019.06.009> PMID: 31326723
43. Darki F, Ferrario A, Rankin J. Hierarchical processing underpins competition in tactile perceptual bistability. *J Comput Neurosci.* 2022;51(3):343–60. <https://doi.org/10.1007/s10827-023-00852-0> PMID: 37204542
44. Snyder JS, Elhilali M. Recent advances in exploring the neural underpinnings of auditory scene perception. *Ann N Y Acad Sci.* 2017;1396(1):39–55. <https://doi.org/10.1111/nyas.13317> PMID: 28199022
45. Rankin J, Sussman E, Rinzel J. Neuromechanistic model of auditory bistability. *PLoS Comput Biol.* 2015;11(11):e1004555. <https://doi.org/10.1371/journal.pcbi.1004555> PMID: 26562507
46. Rankin J, Osborn Popp PJ, Rinzel J. Stimulus pauses and perturbations differentially delay or promote the segregation of auditory objects: Psychoacoustics and modeling. *Front Neurosci.* 2017;11:198. <https://doi.org/10.3389/fnins.2017.00198> PMID: 28473747
47. Ferrario A, Rankin J. Auditory streaming emerges from fast excitation and slow delayed inhibition. *J Math Neurosci.* 2021;11(1):8. <https://doi.org/10.1186/s13408-021-00106-2> PMID: 33939042
48. Moreno-Bote R, Rinzel J, Rubin N. Noise-induced alternations in an attractor network model of perceptual bistability. *J Neurophysiol.* 2007;98(3):1125–39. <https://doi.org/10.1152/jn.00116.2007> PMID: 17615138
49. Shpiro A, Moreno-Bote R, Rubin N, Rinzel J. Balance between noise and adaptation in competition models of perceptual bistability. *J Comput Neurosci.* 2009;27(1):37–54. <https://doi.org/10.1007/s10827-008-0125-3> PMID: 19125318

50. Kondo HM, Pressnitzer D, Shimada Y, Kochiyama T, Kashino M. Inhibition-excitation balance in the parietal cortex modulates volitional control for auditory and visual multistability. *Sci Rep*. 2018;8(1):14548. <https://doi.org/10.1038/s41598-018-32892-3> PMID: 30267021
51. Bregman AS. Auditory streaming: Competition among alternative organizations. *Percept Psychophys*. 1978;23(5):391–8. <https://doi.org/10.3758/bf03204141> PMID: 683823
52. Dancer C. Dancer design company. <https://www.dancerdesign.co.uk/products.html>.
53. Parzen M, Ghosh S, Lipsitz S, Sinha D, Fitzmaurice GM, Mallick BK, et al. A generalized linear mixed model for longitudinal binary data with a marginal logit link function. *Ann Appl Stat*. 2011;5(1):449–67. <https://doi.org/10.1214/10-AOAS390> PMID: 21532998
54. Siegel LK, Silva M, Lin L, Chen Y, Liu Y-L, Chu H. Choice of link functions for generalized linear mixed models in meta-analyses of proportions. *Res Methods Med Health Sci*. 2025;6(1):13–23. <https://doi.org/10.1177/26320843231224808> PMID: 39668973
55. Goulet-Pelletier J-C, Cousineau D. A review of effect sizes and their confidence intervals, Part I: The Cohen's d family. *TQMP*. 2018;14(4):242–65. <https://doi.org/10.20982/tqmp.14.4.p242>
56. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav Res Methods*. 2009;41(4):1149–60. <https://doi.org/10.3758/BRM.41.4.1149> PMID: 19897823
57. Micheyl C, Tian B, Carlyon RP, Rauschecker JP. Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron*. 2005;48(1):139–48. <https://doi.org/10.1016/j.neuron.2005.08.039> PMID: 16202714
58. Langers DRM, van Dijk P. Mapping the tonotopic organization in human auditory cortex with minimally salient acoustic stimulation. *Cereb Cortex*. 2012;22(9):2024–38. <https://doi.org/10.1093/cercor/bhr282> PMID: 21980020
59. Wu GK, Arbuckle R, Liu B-H, Tao HW, Zhang LI. Lateral sharpening of cortical frequency tuning by approximately balanced inhibition. *Neuron*. 2008;58(1):132–43. <https://doi.org/10.1016/j.neuron.2008.01.035> PMID: 18400169
60. Fishman YI, Kim M, Steinschneider M. A crucial test of the population separation model of auditory stream segregation in macaque primary auditory cortex. *J Neurosci*. 2017;37(44):10645–55.
61. Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ. Neuromagnetic correlates of streaming in human auditory cortex. *J Neurosci*. 2005;25(22):5382–8. <https://doi.org/10.1523/JNEUROSCI.0347-05.2005> PMID: 15930387
62. Gutschalk A, Micheyl C, Oxenham AJ. Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biol*. 2008;6(6):e138. <https://doi.org/10.1371/journal.pbio.0060138> PMID: 18547141
63. Dykstra AR, Halgren E, Thesen T, Carlson CE, Doyle W, Madsen JR, et al. Widespread brain areas engaged during a classical auditory streaming task revealed by intracranial EEG. *Front Hum Neurosci*. 2011;5:74. <https://doi.org/10.3389/fnhum.2011.00074> PMID: 21886615
64. Fernandez Pujol C, Blundon EG, Dykstra AR. Laminar specificity of the auditory perceptual awareness negativity: A biophysical modeling study. *PLoS Comput Biol*. 2023;19(6):e1011003. <https://doi.org/10.1371/journal.pcbi.1011003> PMID: 37384802
65. Lankinen K, Ahveninen J, Jas M, Raij T, Ahlfors SP. Neuronal modeling of cross-sensory visual evoked magnetoencephalography responses in the auditory cortex. *J Neurosci*. 2024;44(17):e1119232024. <https://doi.org/10.1523/JNEUROSCI.1119-23.2024> PMID: 38508715