# Multisensory integration enhances audiovisual responses in the Mauthner cell

Santiago Otero Coronel[1,2], Thomas Preuss[3] and Violeta Medan[1,4]*

1. Instituto de Fisiología y Biología Molecular y Celular, Concejo Nacional de Investigaciones Científicas y Tecnológicas.
2. The Rockefeller University, New York, NY, USA.
3. Dept. Psychology, Hunter College, City University of New York, New York, NY, USA.
4. Dept. Fisiología y Biología Molecular y Celular, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.

*Corresponding author.

ABSTRACT

Multisensory integration combines information from multiple sensory modalities to create a coherent perception of the world. In contexts where sensory information is limited or equivocal, it also allows animals to integrate individually ambiguous stimuli into a clearer or more accurate percept and, thus, react with a more adaptive behavioral response. Although responses to multisensory stimuli have been described at the neuronal and behavioral levels, a causal or direct link between these two is still missing. In this study, we studied the integration of audiovisual inputs in the Mauthner cell, a command neuron necessary and sufficient to trigger a stereotypical escape response in fish. We performed intracellular recordings while presenting a diverse range of stimuli to determine which stimulus properties affect their integration. Our results show that stimulus modality, intensity, temporal structure, and interstimulus delay affect input summation. Mechanistically, we found that the distinct decay dynamics of feedforward inhibition triggered by auditory and visual stimuli can account for certain aspects of input integration. Altogether, this is the first study to investigate multisensory integration in a cell of clear behavioral relevance and provides a phenomenological and mechanistic characterization of how multisensory integration depends on stimulus properties.

## INTRODUCTION

The ability to detect a potential threat and react with a startle/escape behavior is critical for survival. However, escapes are metabolically costly and disruptive of other behaviors, therefore tightly controlled by neural networks that only trigger a response when danger is almost certain. Higher activation thresholds in these neural circuits can reduce the rate of false alarms, requiring clear and salient sensory inputs to trigger a behavioral response. In nature, however, many predators develop strategies such as camouflage or partial hiding to decrease their salience, and sensory stimuli tend to be noisy and ambiguous. Consequently, it is adaptive for animals to have ways to disambiguate the signals from the sensory inflow they receive, especially in contexts where available information is limited.

One strategy for reducing sensory uncertainty is multisensory integration (MSI), which is defined as the process of combining information coming from independent sensory streams to form a new percept with reduced ambiguity (Stein and Stanford, 2008). MSI has been extensively documented in animals from nematodes to humans and in a broad range of behavioral contexts. In the context of escape behaviors, MSI of threat-like stimuli increases the escape probability and reduces the latency of the evoked response. In mammals, activation of different parts of the superior colliculus determines whether animals engage in exploratory or defensive behaviors (Dean et al., 1989; Wei et al., 2015). Importantly, the superior colliculus contains neurons that integrate multisensory stimuli due to synaptic convergence of afferents conveying information from different sensory modalities (Kadunce et al., 2001). These studies established that the effects of MSI are larger when weak or ambiguous stimuli are combined, determining an inverse effectiveness relationship with stimulus strength. The underlying hypothesis is that integration of sensory signals from different modalities can only form a more accurate world percept if each unisensory signal carry some uncertainty or ambiguity.

In goldfish, audiovisual MSI increases the probability and reduces the latency of startle escape behaviors, with an inverse relationship between the MSI effectiveness and the stimulus salience (McIntyre and Preuss, 2019; Martorell and Medan, 2022). These startle escapes (C-starts) are triggered by a pair of reticulospinal neurons (Mauthner cells) that receive visual and auditory inputs through anatomically segregated pathways. Auditory inputs reach the M-cell lateral dendrite through a disynaptic (2 ms) pathway, while the visual inputs relay information in the optic tectum before reaching the M-cell ventral dendrite through a polysynaptic (20 ms) pathway (Zottoli et al., 1987; Korn and Faber, 2005; Szabo et al., 2007). The accessibility of M-cells for *in vivo* intracellular recordings and that fact that there is one-to-one relationship between a single M-cell action potential and the initiation of a C-start is well suited to investigate the cellular mechanisms of multisensory integration

with a clear functional consequence (Zottoli, 1977; Weiss et al., 2006). We recently showed that cable properties differ between the auditory and visual M-cell dendrites, putatively to accommodate the different temporal dynamics, excitatory-inhibitory balances, and relative strengths of the respective inputs (Medan et al., 2017). However, it is unclear how auditory and visual signals are integrated at the level of the M-cell to generate a multisensory response. Here we study how visual and auditory presynaptic pathways and intrinsic characteristics of an identified multisensory neuron interact to produce MSI to ultimately enhance threat detection. To this aim we performed somatic intracellular recordings of goldfish M-cell *in vivo* using auditory and tectal stimuli with naturalistic dynamics. Our results show how the temporal dynamics, stimulus strength and feed-forward inhibition affect MSI and reveals an inverse effectiveness relationship between MSI integration and stimulus intensity. This represents a rare characterization of MSI at a single-cell level in a neuron that is sufficient and necessary to trigger a stereotyped behavior.

METHODS

**Animals**

We used adult goldfish (*Carassius auratus*) of both sexes and 7–10 cm of standard body length. Fish were purchased from Billy Bland Fishery (Taylor, AR, USA), Hunting Creek Fisheries (Thurmont, MD, USA) or Daniel Corriarello Aquarium (Buenos Aires, Argentina) and allowed to acclimate for at least a week after transport. Fish were kept in rectangular Plexiglas holding tanks (30 x 60 x 30 $cm^3$; 54 l) in groups of up to 10 individuals. Tanks were supplied with filtered and dechlorinated water and maintained at 18 °C. Ambient light was set to a 12-h light/dark photoperiod. Animals were fed floating pellets (Sera, Germany) five times a week. All procedures and protocols were performed in accordance with the guidelines and regulations of the Institutional Animal Care and Use Committee of Hunter College, City University of New York and Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.

**Acoustic stimuli**

Sound stimuli consisted in single-cycle sound pips (5-ms duration at 200 Hz) produced by a function generator (33210A Agilent Technologies Inc., Santa Clara, CA, USA) connected to a shielded subwoofer (SA-WN250 Sony Corp., Tokyo, Japan) located at 30 cm from the recording chamber. Sound stimuli were recorded with a microphone positioned 10 cm over the head of the fish. For calibration, a hydrophone (SQ01, Sensor Technology, Ontario, Canada) was positioned where the submerged fish was placed during experiments. As a result of transfer loss through the media of the recording chamber (air to water), the maximum underwater sound intensity was 147 dB SPL relative (re) to 1 $\mu$Pa in water. We deliberately used sound intensities which are known to be subthreshold for evoking behavioral startle responses (Neumeister et al., 2008) as we were interested on studying integration of subthreshold stimuli.

**Optic tectum stimulation**

Tectal stimuli consisted in the electrical stimulation of the mid-anterior rim of the left optic tectum using a bipolar electrode (#30202 FHC, Bowdoin, ME, USA) connected to an isolated stimulator (DS2A; Digitimer Ltd, Welwyn Garden City, UK). In most cases, stimulation [100 µs, 0.1-1mA] in this area was effective in driving significant responses (> 4 mV) in the M-cell and we did not change the position of the electrode or the amplitude of the current during the experiment. To change the strength and dynamics of the tectal stimulation we either used a 60-Hz stimulation of variable duration (1, 33, 66, 100, 200 ms) or a 100-ms train of variable frequency (1, 33, 100, 200 Hz).

**Electrophysiology**

Intracellular responses of the M-cell to tectal and acoustic stimuli were studied *in vivo* using standard surgical and electrophysiological recording techniques (Preuss and Faber, 2003; Preuss et al., 2006; Medan et al., 2017). To initiate anesthesia, fish were immersed in 1 liter of ice water with 40 mg/l of the general anesthetic tricaine methanesulfonate (MS-222, Western Chemical, Ferndale, WA, USA), until the fish ceased to swim, lost equilibrium and were unresponsive to a pinch on the tail (typically 10–15 min). They were next treated with 20% benzocaine gel (Ultradent, South Jordan, UT, USA) at incision sites and pin-holding points 5 min prior to surgical procedures. Fish were stabilized in the recording chamber by two pins, one on each side of the head, and ventilated through the mouth with recirculating, aerated saline at 18°C (saline [g/l]: sodium chloride 7.25, potassium chloride 0.38, monosodium phosphate monobasic 0.39, magnesium sulphate 0.11, Hepes 4.77; calcium chloride 0.24; dextrose 1.01, pH 7.2). The recording chamber was mounted inside an opaque, thin-walled tank filled with saline that covered the fish up to eye level. The recirculating saline also included a maintenance concentration of the anesthetic MS-222 (20 g/l) that does not interfere with auditory processing (Palmer and Mensinger, 2004; Cordova and Braun, 2007). Next, the spinal cord was exposed with a small lateral incision at the caudal midbody. Bipolar stimulation electrodes were placed on the unopened spinal cord to transmit low-intensity (5–8 V) electrical pulses generated by an isolated stimulator (A 360, WPI, Sarasota, FL, USA). This allowed antidromic activation of the M-cell axons, as confirmed by a visible muscular contraction (twitch). Surgical procedures were performed before a muscle paralysis agent was injected, which allows monitoring of the effectiveness the anesthetic by watching for an increase of opercula movement frequency (largely reduced in deep anesthesia) and movements/twitches in response to the surgical procedures. Shortly before the recordings started, animals were injected I.M. with D-tubocurarine (1 $\mu$g g−1 b.w.; Abbott Laboratories, Abbott Park, IL, USA) and a small craniotomy exposed the medulla for electrophysiological recordings. Antidromic stimulation produces a negative potential in the M-cell axon cap (typically 15–20 mV), which unambiguously identifies the axon hillock and allows intracellular recordings from defined locations along the M-cell soma-dendritic membrane (Furshpan and Furukawa, 1962; Furukawa, 1966; Faber and Korn, 1989). Intracellular recordings were acquired using borosilicate glass electrodes (7–10 M$\Omega$) filled with 5 M potassium acetate and an Axoclamp-2B amplifier (Axon Instruments, Foster City, CA, USA) in current-clamp setting. M-cell responses were acquired with a Digidata 1440A (Axon Instruments) at 25 kHz. Electrodes were advanced using motorized micromanipulators (MP-285; Sutter Instruments, Novato, CA, USA) until reaching the axon cap (defined as a site with an extracellular M-cell AP field >10 mV). Next, the electrode was moved 50

$\mu$m lateral and 50 $\mu$m posterior to penetrate the somatic region. Only trials in which the resting membrane potential was between -90 and -70 mV were included in the analysis.

**Data Analysis**

Postsynaptic responses to acoustic and tectal or visual stimuli were recorded at the M-cell soma and averages of 5-10 responses were used in the statistical analysis. We quantified responses calculating the mean evoked depolarization in a 12 ms window after pip onset (auditory only trials, A) or after the last tectal stimulus in tectal-only (tectal only trials, T) and multisensory (tectal + auditory trials, M) trials (Supp. Fig. 1A-B). We also measured the peak amplitude of the response, which in nearly-all cases occurred in the same time window. As peak amplitude analysis produced the same qualitative results, only the mean area analysis is shown.

To characterize the multisensory integration, the responses to M stimuli were normalized to the T and A components. In the field of multisensory integration, multisensory responses are typically normalized to the maximum response of the unisensory (Meredith and Stein, 1983; Stanford and Stein, 2007; Stein and Stanford, 2008), which we call Multisensory-Maximum Unimodal index (MSI/Max):

$$MSI/Max = M \div \max(T, A)$$

The value of this index determines the net effect of the integration, with MSI/Max greater than 1 representing multisensory enhancement, MSI/Max lower than 1 representing multisensory inhibition, and MSI/Max not significantly different than 1 suggesting absence of multisensory integration (Stein and Stanford 2008). Additionally, to calculate the linearity of the multisensory integration performed by the M-cell, we calculated the ratio between the multisensory response and the sum of the unisensory responses, which we call Multisensory-Sum Unimodals index (MSI/Sum):

$$\mathrm{MSI/Sum} = M \div (T + A)$$

The value of this index determines the linearity of the integration, with $\mathrm{MSI/Sum}$ greater than 1 representing supralinear integration, $\mathrm{MSI/Sum}$ of 1 representing linear integration, and $\mathrm{MSI/Sum}$ lower than 1 representing sublinear integration (Stanford et al., 2005).

One of the principles that guide multisensory integration is the spatial coherence of the different cues being integrated. If the perceived source localization of two stimuli overlap, these stimuli are more likely be integrated. Conversely, if the two stimuli arrive from spatially disparate sources the chances of observing a multisensory integration will decrease (Meredith and Stein, 1996; Stein and Stanford,

2008). The geometrical configuration of the setup in our experiments determined that the loudspeaker used to produce acoustic stimuli was mounted to the right of the fish and electrical stimulation was applied to the left tectum (which normally receives visual signals from the right retina). Whenever possible, we performed recordings from both Mauthner cells to minimize the number of fish used in the study. To rule out the possibility of a systematic difference on the responses of either M-cell due to the spatial source of the stimuli, we compared the responses to tectal or auditory stimuli in sequential recordings of the right and left M-cells of the same fish. Figure Supplementary 1 shows representative trace examples for tectal (100 ms 60 Hz tectal train, Fig. Supp. 1A) and auditory (5 ms 200 Hz auditory pip, Fig. Supp. 1B) stimulation from the left (upper traces) or right (lower traces) M-cell of the same fish. The complex post-synaptic potentials were highly stereotypical in successive presentations of a T (or A) stimulus within and across cells. Analysis from pairs of tectal (Supp. Fig. 1C, N=9) and auditory (Supp. Fig. 1D, N=11) M-cell responses showed no statistical differences between ipsi- and contralateral stimulation (paired T-test, tectal: p=0.788; auditory: p=0.467). This confirms previous results indicating that in goldfish, tectal stimulation produces responses with similar latency and amplitude in both M-cells (Zottoli et al., 1987; Canfield, 2006). We therefore pooled data acquired from left or right M-cells for all following analysis.

**Statistics**

To perform statistical analysis we used R (version 4.0.2, www.r-project.org) and RStudio (version 1.1.456, www. rstudio. com). A significance level of $\alpha$ = 0.05 was used throughout the study. Effects of explanatory variables over response variables were assessed using Linear Models (binomial GLMs in the case of binary variables, Linear Models (LM, followed by ANOVA) in the case of continuous variables). Paired T-tests were used to compare tectal or auditory responses recorded in left or right M-cells (see Fig. Supp. 1). Wilcoxon Ranked Sum test was used for single sample comparisons. Sample size is denoted by *N* when it refers to the number of M-cells or *n* when it refers to the number of trials (average of 5-10 presentations). Boxplots show the median, 25 and 75$^{th}$ percentile, and lines indicate minimum and maximum values.

RESULTS

## MAUTHNER CELL RESPONSES ARE SENSITIVE TO THE DYNAMICS OF VISUAL TECTAL STIMULI

Multisensory integration is known to be dependent on the saliency of the unimodal components (Stanford et al., 2005; Martorell and Medan, 2022). In order to create a stimulus set of variable strength and temporal dynamics, we recorded M-cell responses to different electrical trains delivered in the optic tectum. M-cell responses to tectal stimulation are reminiscent of the responses of naturalistic visual inputs such as looming stimuli (Preuss et al., 2006; Medan and Preuss, 2014). However, responses to visual looms are significantly affected by repeated stimulation and anesthesia while responses to tectal trains remained stable for the duration of the experiments.

M-cells respond to tectal trains of electrical stimulation with phasic postsynaptic potentials riding on top of a tonic depolarizing response (Figure 1A') that reaches a plateau after 30-100 ms and decays to resting values 200-300 ms after the end of the tectal train. However, it is not known how the frequency and duration of the stimulation trains in the optical tectum affect the magnitude and dynamics of the M-cell responses. To characterize how the M-cell responses depend on the intensity and structure of the tectal stimulation, we recorded responses to 60-Hz trains that varied in duration (Figure 1A-B), and 100-ms tectal trains that varied in pulse frequency (Figure 1C-D). The responses to both sets of trains exhibited complex temporal dynamics with a tonic component that progressively builds during the train and faster phasic components after each pulse.

Increasing the duration of the 60-Hz trains drove stronger tectal responses, due to the combined effect of increasing both the tonic and phasic components of the response ($p<0.0001$, Figure 1A, B). Similarly, increasing the frequency of the 100-Hz stimulation trains also drove stronger tectal responses, due to a large increase in the tonic component and a modest but still significant increase in the phasic component ($p<0.0001$ and $p=0.02$ respectively, Figure 1D, E). Comparing stimulus trains of variable duration and frequency, modifying the duration of the 60-Hz trains was more effective in tuning the magnitude of the responses (compare the total response amplitude -phasic + tonic- for 1 and 200 ms) than the 100-ms trains (compare 30 vs. 200 Hz). Since multimodal integration is known to depend on the strength of the unimodal components (Stanford et al., 2005; Martorell and Medan, 2022), in the following multimodal experiments we leveraged this characterization to generate tectal stimuli of variable strength and dynamics. We decided to include both the 60-Hz trains of variable duration and the 100-ms trains of variable frequency since this would allow us to determine whether any effect was due to the magnitude or structure (frequency and duration) of the stimulus.

## MULTISENSORY ENHANCEMENT IN THE MAUTHNER CELL

Although the M-cell is known to receive inputs and respond to stimuli of different modalities, there are no reports on how it combines and integrates information when two stimuli of different modalities are co-presented. To determine whether multisensory enhancement occurs in the M-cell, we recorded somatic responses to an auditory pip (unisensory auditory, A), a tectal stimulation train (unisensory tectal, T), and their combination (multisensory audiotectal, M) (Figure 2A-B). Importantly, since we were interested in whether and how the integration of a multisensory stimulus is affected by the intensity and dynamics of its unisensory components, we varied the duration and frequency of the tectal stimulation for the T and M stimuli.

Increasing the duration of the 60-Hz tectal train in an M stimulus positively modulated the multisensory responses (amplitudes ranging from 3.10 ± 0.95 mV to 5.43 ± 1.80 mV, N = 17, n =74; LM: evoked depolarization ~ tectal duration, $p<0.001$, Figure 2B). We then compared the multisensory responses to their unisensory counterparts. When M responses were normalized to the most effective unisensory response, all multisensory stimuli yielded MSI/Max indices that ranged from 1.43 to 1.54 for all stimulus durations (Wilcoxon single sample test mu=1, $p<0.001$ for all stimulus durations, Figure 2C). When the M response was normalized to the sum of both unisensory responses the MSI/Sum indices were lower than 1 (i.e. sublinear integration, Wilcoxon single sample test mu=1, alternative mu<1, $p<0.001$) for all cases except for the briefest tectal stimulus, for which the multisensory integration was supralinear (Wilcoxon single sample test mu=1, alternative mu>1, $p<0.01$, Figure 2D). Increasing the frequency of a 100-ms tectal train (Figure 2E) was not as effective in increasing the multisensory responses to M (LM: evoked depolarization ~ tectal frequency, $p=0.07$, Figure 2F). However, the responses to these M still exhibited MSI/Max indices that ranged 1.44-1.53 when compared to the largest unimodal stimulus (Figure 2G). The MSI/Sum of these M responses ranged 0.82 to 0.86 (Wilcoxon single sample test mu=1, alternative mu<1, all $p<0.001$, Figure 2H). Neither MSI/Max o MSI/Sum showed a significant dependency with tectal frequency (LM: MSI/Max ~ tectal frequency, $p=0.98$, MSI/Sum ~ tectal frequency, $p=0.27$).

Overall, these results show that multisensory integration in the M-cell enhances sensory responses by about 50% of the most effective unisensory stimulus, regardless of the duration and frequency of the tectal stimulus. While most multisensory stimuli exhibited a sublinear integration, the multisensory stimulus with the shortest (and weakest) tectal component produced a slightly but significantly supralinear response.

**TECTAL AND AUDITORY-EVOKED FEEDFORWARD INHIBITIONS EXHIBIT DISTINCT TEMPORAL DYNAMICS**

Both visual and auditory stimuli evoke feedforward inhibition (FFI) at the perisomatic region of the M-cell through a population of inhibitory interneurons open chloride channels that produce shunting of the postsynaptic currents (Diamond et al., 1973; Zottoli et al., 1987; Koyama et al., 2011; Tabor et al., 2018). The functional roles of FFI include restricting the temporal window of sensory excitation and terminate the excitatory signal. In combination with passive membrane decay, an increase in FFI will further narrow the window for temporal integration Conversely, a decrease in FFI would increase lingering of a membrane depolarization produced by a previous stimulus, and thus extend the temporal integration window Thus differences in the strength or temporal dynamics between visual and auditory evoked FFI inhibition may affect multisensory integration. In the M-cell, auditory or tectal evoked FFI can be quantified as the amplitude reduction of an antidromically stimulated "test" action potential (APtest) (Faber and Korn, 1982, Medan et al. 2017). We used "conditioning" stimuli of similar strength specifically, a single auditory pip and a brief tectal train (~5 mV) evoking comparable depolarizations in the M-cell. To test if FFI evoked by auditory or tectal stimuli differed, we evoked the APtest at specific times (2–80 ms) following a sensory stimulus and calculate shunting inhibition as %SI= 100 - APtest/APcontrol ×100, where AP control is the amplitude of the AP without a conditioning stimulus (Figure 3A). FFI inhibition induced by auditory and tectal stimulation showed similar onset latency and maximum amplitude (Figure 3B, auditory vs. tectal %SI: 13.65 ± 4.74 vs. 14.55 ± 7.50; Kruskal-Wallis chi-squared = 0.01, p = 0.91) with peak inhibition at the same time (auditory vs. tectal time (ms) of maximum %SI: 9.04 ± 3.93 vs. 8.85 ± 6.54; Kruskal-Wallis chi-squared = 0.30, p = 0.58, Figure 3C). In a subset of experiments (N = 7), both auditory and tectal FFI were evaluated sequentially for each cell and same recording site which allowed us to have an accurate estimate of the inhibition triggered by each sensory modality (Figure 3D). Although the maximum level of inhibition was comparable and was reached in a similar time, recovery from inhibition was much faster for tectal than for auditory FFI. Tectal FFI decayed to 50% of its peak in 8 ms, compared to 30 ms for the auditory FFI (Kruskal-Wallis chi-squared = 5.97, p = 0.01, Figure 3E). The area under the curve for auditory FFI was almost 3 times greater than for tectal FFI (502 vs. 176 %.ms, respectively, Kruskal-Wallis chi-squared = 5.59, p = 0.02, Figure 3F).

These results further support the key role of the temporal dynamics and the modality of the stimuli that are integrated in the M-cell. The FFI affects the decay rate of the excitatory response and effectively modulates to which extent and the temporal windows in which different stimuli could be integrated. The short-lived FFI that follows tectal stimulation suggests that multiple trains or waves of tectal inputs are more likely to be integrated and summed than multiple sounds. This goes in line with duration of effective visual looms (in the order of hundreds of ms) (Preuss et al., 2006; Temizer et al.,

2015) and the sounds (in the order on tens of ms) (Casagrand et al., 1999; Weiss et al., 2009) that are most effective in triggering the M-cell and eliciting C-start escapes.

**STIMULUS TEMPORAL ORDER AND MODALITY AFFECT STIMULI INTEGRATION**

There are several anatomical and functional differences in the auditory and visual inputs that eventually reach the M-cell soma. First, auditory inputs reach the lateral dendrite of the M-cell through a disynaptic (≈2ms) pathway, while visual inputs reach the ventral dendrite after a polysynaptic (≈20ms) pathway that includes the optical tectum (Zottoli et al., 1987; Szabo et al., 2006). Second, M-cell dendritic characteristics determine modality specific filtering differences resulting in stronger attenuation of visual than auditory signals (Medan et al., 2017). For auditory stimuli, it has been reported that a first sound reduces the M-cell response to a second sound, a phenomenon known as auditory prepulse inhibition (Neumeister et al., 2008; Curtin and Preuss, 2015; Tabor et al., 2018). However, it is unclear whether a similar phenomenon would hold true for cross-modal and tectal-tectal interactions, since auditory and visual inputs travel through distinct and separated pathways and undergo different filtering.

To investigate the role of input modality on M-cell integration we used pairs of single auditory pips and single tectal pulses 50 ms apart with in all possible sequences, including unisensory combinations: AT, TA, AA and TT (Figure 4A). The amplitude of the tectal stimulation pulse was chosen so that the M-cell responses to the tectal and auditory stimuli were of similar magnitude (4-8 mV). These initial responses decayed passively over the course of the 50 ms following the presentation of the first stimulus. The mean lingering depolarization from the evoked auditory pip in the AA and AT trials (shaded rectangles in Figure 4A) was higher than for the tectal stimulus in the TA and TT trials (1.12 ± 0.56mV vs. 0.75 ± 0.36mV, t-test: $p=0.03$, Figure 4B). This is consistent with previous results showing that visual (tectal) input is subject to stronger decay than auditory inputs (Medan et al., 2017). The response to the second stimulus was calculated using the voltage immediately before its presentation as baseline, which also represented the lingering depolarization. In the AA trials, we observed the previously described prepulse inhibition phenomenon, with the first auditory pip reducing the amplitude of the second pip by 40. To quantify the effect of the first stimulus on the second stimulus and compare the effect of its modality, we calculated the ratio between the amplitude of the second stimulus (A o or T) and the lingering depolarization plus the amplitude without a preceding stimulus (i.e. in a “unisensory condition”). Notably, the inhibitory effect observed in the AA trials was not present for the other cross-modal and tectal-tectal combinations, with all of them showing no reduction in the amplitude of the response to the second stimulus (Figure 4C, LM, $F=72.7$, $p<0.001$).

**MULTISENSORY INTEGRATION IN THE MAUTHNER CELL SHOWS INVERSE EFFECTIVENESS**

The principle of inverse effectiveness states that the effects of multisensory integration are typically larger when the components of a multisensory stimulus are weak. Throughout our experiments, we observed that the M-cell only integrated supralinearly when presenting M stimuli with short tectal trains (Figure 2C, 1ms tectal), linearly when presenting AT – TA, and sublinearly for M stimuli with longer tectal trains (Figure 2D-H). However, these results represent the average MSI for each group for a specific set of experimental conditions and effectively simplify the diverse range of responses across trials, experimental conditions, and fish. To compare the overall dependence of the multisensory enhancement on the strength of sensory inputs, we pooled the multisensory integration indexes for all fish and for all experiments irrespective of sequence of stimuli, tectal duration and frequency (Figure 5A, MSI/Sum and B, MSI/Max). We found a significant inverse correlation between the magnitude of the multisensory enhancement and the magnitude of the unisensory components (LM: MSI/Sum, $p<0.001$; MSI/Max, $p<0.001$). This result parallels the inverse effectiveness reported at the behavioral level in goldfish (McIntyre and Preuss, 2019; Martorell and Medan, 2022), confirming that the relative multisensory enhancement obtained by combining two inputs is proportionally higher when the stimuli are weak. Combinations of weak stimuli produce supralinear multisensory responses that are between 50-100% stronger than the most effective unisensory stimulus (Figure 5A-B). Noteworthy, these results provide robust evidence that in the escape network of fish, inverse effectiveness of multisensory integration is implemented, at least, at the single (Mauthner) cell level. Although our analysis was based on subthreshold responses, in some cases we observed that the multisensory enhancement of the response was sufficient for surpassing the threshold and generated action potentials in the M-cell. Figure 5C shows two examples from two different animals in which the M-cells showed subthreshold responses to unisensory stimuli and elicited an action potential when the two stimuli are combined. Since an action potential in the M-cell triggers the initiation of a C-start escape, these are examples of behaviors that only occurred in response to a multisensory stimulus but not to either of the unisensory components. We did not see any action potential in response to unisensory stimuli across 18 animals. This highlights how, because of the all-or-nothing nature of action potentials, relatively small increments in the M-cell response due to multisensory enhancement could be sufficient to make the cell trigger and have a profound effect in escape behavior, even if the integration is sublinear.

DISCUSSION

The Mauthner-cell system triggers a behavior of vital importance: escaping from predators. A combination of fast afferent circuits, synaptic specialization, special intrinsic properties and topology of the M-cell contribute in concert to a fast startle escape response (Korn and Faber, 2005, Eaton et al. review). Here we show that the M-cell incorporates multisensory information to enhance its responses and characterize how the integration of multiple stimuli is affected by the modality, spatial source, temporal structure, temporal order, and intensity of the stimuli. Although multisensory enhancement has been previously described in different neural populations and across a wide range of species, it was unclear whether the activity of these neurons was necessary, sufficient, or even influenced the execution of behavioral responses to multisensory stimuli. In contrast to other known multisensory structures of vertebrates multisensory integration in the M-cell has an unequivocal function in fish behavior and survival, as it directly determines whether and when a C-start is triggered (McIntyre and Preuss, 2019; Martorell and Medan, 2022) .

Different factors differentially affect the summation and integration of auditory and visual inputs to the M-cell. First, sensory inputs are subject to different decays depending on the site of the dendritic arbor at which they arrive. Auditory inputs reach the M-cell lateral dendrite through dense mixed electrical and chemical synapses driven through the eighth nerve afferents (Szabo et al., 2006). In contrast, visual inputs arriving to the ventral dendrite establish synapses that do not show additional specializations as the club endings (Zottoli et al., 1987; Cachope et al., 2007; Szabo et al., 2007). Second, the geometrical properties of the M-cell lateral and ventral dendrites (length and tapering) impose different filters on the input signal which determine a stronger spatial decay for visual than for auditory signals (Medan et al., 2017). Third, as excitatory signals reach the soma, the active ion conductances decrease the membrane resistance and impact the response to following signals (Hodgkin, 1947).

These factors modulate the effective multisensory integration in the M-cell soma, which emerges as a complex, nonlinear interaction between “positive” (additional input strength and summation) and “negative” (temporal and spatial decay, feedforward inhibition and depolarization-induced decrease in input resistance) contributions. Although these factors likely influence integration in any sensory neuron, the accumulated knowledge on the presynaptic elements and the intrinsic mechanisms operating in the M-cell combined with its functional role in commanding the initiation of the C-start response makes this neuron a unique model for analyzing multisensory integration in a single identifiable neuron.

We have previously shown that fish exhibit a behavioral multisensory enhancement of the C-start escape response for audiovisual stimuli, and that this effect is reduced when salience of the unisensory

stimuli increases (Martorell and Medan, 2022). Therefore, one of our goals here was to investigate if the inverse effectiveness principle observed behaviorally had its correlate in integration processes taking place in the M-cell. Paralleling our previous behavioral results, here we show through intracellular recordings that, the response evoked by an auditory pip can be increased by about 50% when it is preceded by a tectal stimulus (Figure 2C, G). Moreover, we found that integration in the M-cell shows inverse effectiveness with sensory input strength (Figure 4 A-B). However, for most multisensory stimuli, the response is smaller than the sum of the evoked depolarization of the two components (Figures 2D, H). This sublinear integration was expected, since cellular biophysics predict a sublinear integration of subthreshold inputs due to a reduction in the driving force of sodium and the aperture of ion channels that reduce the membrane resistance. Noteworthy, action potentials are all-or-nothing events, so the effects when the membrane potential is close to the threshold, sublinear summation can determine that an otherwise silent neuron fires an AP (Figure 5C). Therefore a weak (subthreshold) auditory pip combined by a weak visual stimulus in close temporal association might be sufficient for triggering an escape response.

Importantly, for the weakest multisensory stimuli the M-cell response deviated from this sublinear integration (Figures 2C and 5A). When the sum of unisensory components is less than 2.5 mV (essentially a single tectal stimulus immediately followed by a single auditory pip) summation is supralinear (Figure 5A). This suggests the existence of one or more active (supralinear) mechanism/s of multisensory integration upstream of the M-cell soma, potentially through a presynaptic neuron that also receives audiovisual inputs and projects to the M-cell or locally at the level of the dendritic tree of the M-cell. Indeed, modelling experiments have suggested that active conductances expressed in the proximal dendrites of the M-cell confer nonlinear properties to the lateral and ventral dendrites that increases their excitability. This in turn could particularly enhance multisensory integration of weak stimuli (Medan et al., 2017).

The temporal structure of the multisensory stimuli is crucial to allow the animal to effectively “bind” or associate both signals as coming from a single source (Stein and Stanford, 2008). In the M-cell, this translates into the time window where the effect on the membrane voltage (excitatory and/or inhibitory) persists. Here, we found that a delay of 50 ms between the components of a low-amplitude multisensory stimuli result in a linear summation of the signals (Figure 4C) and that the temporal order of the components does not affect the integration (Figure 4C, AT vs. TA) (Felch et al., 2016). On the other hand, the integration of unisensory stimuli was dependent on the modality of the stimuli with two tectal (TT) stimuli summing linearly and two auditory (AA) stimuli summing sublinearly (Figure 4C). This difference is probably due to differences in the temporal decay of auditory and tectally evoked FFI (Figure 3D).

Auditory stimuli produce not only a longer-lasting inhibition than tectal stimuli of similar strength, but also generate a slow-decaying depolarization upon which the tectal stimulus can sum. While tectal inhibition is almost extinct 50 ms after a tectal stimulus, auditory inhibition stabilizes at around 50% of the maximum inhibitory effect in a time window of 30 to 70 ms after an auditory stimulus (Figure 4E). Interestingly, the integration of AT stimuli was similar to that of TA, supporting a differential cross-modal and intra-modal integration of the excitatory or inhibitory components. Another factor that could be contributing to the differential integration between AA and AT stimuli is that FFI inhibition is likely to be heterogeneously spread in the perisomatic area around the M-cell. If auditory stimulation recruits inhibitory PHP neurons that target the proximal lateral dendrite of the M-cell, a subsequent sensory input propagating along the lateral dendrite (i.e, auditory input) would suffer stronger attenuation than one reaching the soma through the ventral dendrite (i.e., tectal input). However, specific studies are required to test this hypothesis as it is still unclear if auditory and tectal afferent fibers contact the same or different sets of inhibitory neurons and which is the spatial distribution of those inhibitory contacts. Additionally, presynaptic effects could also contribute to the asymmetry in unimodal auditory (AA) vs. unimodal tectal (TT) o multisensory processing. A subset of glutamatergic Gsx1 neurons located in rhombomere 4 have been shown to mediate M-cell auditory presynaptic inhibition (Tabor et al., 2018). These neurons produce presynaptic inhibition of $8^{th}$ nerve auditory fibers adjacent to the distal portion of the M-cell lateral dendrite but leave unaffected the ventral dendrite. This mechanism could therefore affect the processing of pairs of auditory signals (AA) and leave the integration of tectal or multisensory stimuli unaffected.

FIGURE LEGENDS

**Supplementary Figure 1**. A-B. Representative recordings of left or right M-cell somatic responses to an electric stimulus train delivered to the left tectum (A) or to a short 200 Hz sound pip produced by a loudspeaker placed to the right side of the fish (B). The mean (of 3-5 repetitions) response is shown in solid color and either the diagram of the tectal stimulus or the microphone recording of the sound pip are shown below the traces. Dashed boxes indicate the 12-ms time windows following the stimulus where the responses were quantified. C-D. Boxplots showing the median, upper and lower quartiles and comparing the response to the same stimulus in the left or right M-cell in 10 fish for the tectal stimulus (C) and 11 fish for the auditory stimulus (D). Markers with the same color in left and right belong to pairs of M-cells of the same fish. No differences were found in the evoked response amplitude between the left and right M-cells or for tectal or auditory stimuli.

**Figure 1**. M-cell responses to tectal stimulation. A. M-cell responses to a 60 Hz tectal train of 200 ms were used to measure the evoked depolarization at 1, 33, 100 and 200 ms after stimulus onset during a 12-ms time window (colored vertical boxes). A`, Lower inset. Dark grey and light grey areas in the trace show amplitudes of the tonic and phasic components of the response, respectively. The horizontal dotted line indicates the baseline before stimulation. B. Mean amplitude of the phasic (shaded) and tonic (solid) components of tectal trains for the time windows described in A. Bars indicate ± SEM, N=13-14. C. M-cell responses to a 100 ms tectal pulse of 30, 60, 100 or 200 Hz were recorded and quantified as in A. D. Mean amplitude of the phasic (shaded) and tonic (solid) components of 100-ms tectal trains of different frequency. The bars indicate ± SEM, N=10-12. Stimuli of higher frequency produce a larger depolarization.

**Figure 2**. Multisensory integration in the M-cell. A. Diagram of stimulation (left) and response examples (right). The response examples correspond to a single acoustic pulse (blue), to a tectal train of 200 ms at 60 Hz (red), and to their multisensory combination (purple). B. Stacked bars comparing mean (± SEM, N= 14-23) depolarization evoked by tectal, auditory or multisensory stimuli with tectal stimulus trains of indicated duration. C. Boxplots showing the ratios between the multisensory responses and the maximum unisensory response (MSI/Max). D. Boxplots showing the ratios between the multisensory responses and the sum of its unisensory components (MSI/Sum). E-H. As in A-D but for 100 ms tectal trains of different frequencies (N= 14-23) .

**Figure 3**. Auditory and tectal FFI have different decay dynamics. A. Diagram of auditory or tectal evoked feedforward inhibition circuits (center), and the quantification of sensory evoked shunting

inhibition as a reduction of M-cell AP amplitude (auditory: blue, tectal: red; shaded area on top of trace examples) to a control AP (black trace). B-C. Boxplots of peak inhibition and time of peak inhibition triggered by acoustic (N= 21) or tectal stimuli (N= 20). D. Mean (±SEM, N= 7) FFI triggered by tectal and auditory inputs measured in the same M-cells at different time points between 0 and 70 ms post-stimulus. E. Boxplot of the time elapsed to 50% decay of peak FFI for the auditory (left) and tectal (right) stimuli. *p=0.01. F. Area below the curve (0-70 ms) for the auditory (left) and tectal (right) stimuli. *p=0.02.

**Figure 4**. Effect of stimulus sequence and modality on M-cell integration. A. Representative responses of one M-cell to a sequence of acoustic and tectal stimuli, with schematics below each trace showing the modality (auditory: blue, tectal: red) and temporal order of the stimulation. The two stimuli were presented at a 50 ms interval. The shaded blue or red rectangles represent the time windows at which the lingering depolarization of the first responses were measured. B. Lingering depolarization after a single auditory pip or a short tectal train. *p=0.03 (N=29-13). C. Integration in the M-cell for pairs of uni- or multisensory stimuli with a 50 ms delay was calculated as the ratio between the amplitude of the 2$^{nd}$ stimulus and the sum of the lingering depolarization from the response to the first stimulus and the amplitude of the response to the 2$^{nd}$ stimulus without a preceding stimulus (*p<0.001 N=8-29).

**Figure 5**. Multisensory integration of subthreshold unisensory stimuli. A. Pooled responses to different multisensory stimuli, irrespective of duration, frequency, or sequence of stimulation, plotted based on the sum of the responses to its unisensory components and the MSI/Sum index. B. Similar to D, but with the MSI/Max index instead. C. Example traces in which the integration of subthreshold unisensory tectal (T, red) and auditory (A, blue) stimuli evokes an M-cell AP (TA, purple). Responses to an auditory pip only triggered an action potential when they were preceded by four tectal pulses delivered at 333Hz (left) or a 200-ms tectal train at 60Hz (right).

**References**

Cachope R, Mackie K, Triller A, O'Brien J, Pereda AE (2007) Potentiation of electrical and chemical synaptic transmission mediated by endocannabinoids. Neuron 56:1034–1047.

Canfield JG (2006) Functional evidence for visuospatial coding in the Mauthner neuron. Brain Behav Evol 67:188–202.

Casagrand JL, Guzik AL, Eaton RC (1999) Mauthner and reticulospinal responses to the onset of acoustic pressure and acceleration stimuli. J Neurophysiol 82:1422–1437.

Cordova MS, Braun CB (2007) The use of anesthesia during evoked potential audiometry in goldfish (Carassius auratus). Brain Res 1153:78–83.

Curtin PCP, Preuss T (2015) Glycine and GABAA receptors mediate tonic and phasic inhibitory processes that contribute to prepulse inhibition in the goldfish startle network. Front Neural Circuits 9:12.

Dean P, Redgrave P, Westby GWM (1989) Event or emergency? Two response systems in the mammalian superior colliculus. Trends Neurosci 12:137–147.

Diamond J, Roper S, Yasargil GM (1973) The membrane effects, and sensitivity to strychnine, of neural inhibition of the Mauthner cell, and its inhibition by glycine and GABA. J Physiol 232:87–111.

Faber DS, Korn H (1982) Transmission at a central inhibitory synapse. I. Magnitude of unitary postsynaptic conductance change and kinetics of channel activation. J Neurophysiol 48:654–678.

Faber DS, Korn H (1989) Electrical field effects: their relevance in central neural networks. Physiol Rev 69:821–863.

Felch DL, Khakhalin AS, Aizenman CD (2016) Multisensory integration in the developing tectum is constrained by the balance of excitation and inhibition. eLife 5:e15600.

Furshpan EJ, Furukawa T (1962) Intracellular and extracellular responses of the several regions of the Mauthner cell of the goldfish. J Neurophysiol 25:732–771.

Furukawa T (1966) Synaptic Interaction at the Mauthner Cell of Goldfish. In: Progress in Brain Research (Tokizane T, Schadé JP, eds), pp 44–70 Correlative Neurosciences. Elsevier. Available at: https://www.sciencedirect.com/science/article/pii/S0079612308629714 [Accessed October 24, 2022].

Hodgkin AL (1947) The membrane resistance of a non-medullated nerve fibre. J Physiol 106:305–318.

Kadunce DC, Vaughan JW, Wallace MT, Stein BE (2001) The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. Exp Brain Res 139:303–310.

Korn H, Faber DS (2005) The Mauthner cell half a century later: a neurobiological model for decision-making? Neuron 47:13–28.

Koyama M, Kinkhabwala A, Satou C, Higashijima S, Fetcho J (2011) Mapping a sensory-motor network onto a structural and functional ground plan in the hindbrain. Proc Natl Acad Sci U S A 108:1170–1175.

Martorell N, Medan V (2021) Audiovisual integration in the Mauthner cell enhances escape probability and reduces response latency. Available at: https://www.biorxiv.org/content/10.1101/2021.08.19.456957v1 [Accessed August 27, 2021].

Martorell N, Medan V (2022) Audiovisual integration in the Mauthner cell enhances escape probability and reduces response latency. Sci Rep 12:1097.

McIntyre C, Preuss T (2019) Influence of Stimulus Intensity on Multimodal Integration in the Startle Escape System of Goldfish. Front Neural Circuits 13:7.

Medan V, Mäki-Marttunen T, Sztarker J, Preuss T (2017) Differential processing in modality-specific Mauthner cell dendrites. J Physiol.

Medan V, Preuss T (2014) The Mauthner-cell circuit of fish as a model system for startle plasticity. J Physiol-Paris 108:129–140.

Meredith MA, Stein BE (1983) Interactions among converging sensory inputs in the superior colliculus. Science 221:389–391.

Meredith MA, Stein BE (1996) Spatial determinants of multisensory integration in cat superior colliculus neurons. J Neurophysiol 75:1843–1857.

Neumeister H, Szabo TM, Preuss T (2008) Behavioral and physiological characterization of sensorimotor gating in the goldfish startle response. J Neurophysiol 99:1493–1502.

Palmer LM, Mensinger AF (2004) Effect of the anesthetic tricaine (MS-222) on nerve activity in the anterior lateral line of the oyster toadfish, Opsanus tau. J Neurophysiol 92:1034–1041.

Preuss T, Faber DS (2003) Central cellular mechanisms underlying temperature-dependent changes in the goldfish startle-escape behavior. J Neurosci 23:5617–5626.

Preuss T, Osei-Bonsu PE, Weiss SA, Wang C, Faber DS (2006) Neural representation of object approach in a decision-making motor circuit. J Neurosci Off J Soc Neurosci 26:3454–3464.

Stanford TR, Quessy S, Stein BE (2005) Evaluating the Operations Underlying Multisensory Integration in the Cat Superior Colliculus. J Neurosci 25:6499–6508.

Stanford TR, Stein BE (2007) Superadditivity in multisensory integration: putting the computation in context. Neuroreport 18:787–792.

Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. Nat Rev Neurosci 9:255–266.

Szabo TM, McCormick CA, Faber DS (2007) Otolith endorgan input to the Mauthner neuron in the goldfish. J Comp Neurol 505:511–525.

Szabo TM, Weiss SA, Faber DS, Preuss T (2006) Representation of auditory signals in the M-cell: role of electrical synapses. J Neurophysiol 95:2617–2629.

Tabor KM, Smith TS, Brown M, Bergeron SA, Briggman KL, Burgess HA (2018) Presynaptic Inhibition Selectively Gates Auditory Transmission to the Brainstem Startle Circuit. Curr Biol CB 28:2527-2535.e8.

Temizer I, Donovan JC, Baier H, Semmelhack JL (2015) A Visual Pathway for Looming-Evoked Escape in Larval Zebrafish. Curr Biol CB 25:1823–1834.

Wei P, Liu N, Zhang Z, Liu X, Tang Y, He X, Wu B, Zhou Z, Liu Y, Li J, Zhang Y, Zhou X, Xu L, Chen L, Bi G, Hu X, Xu F, Wang L (2015) Processing of visually evoked innate fear by a non-canonical thalamic pathway. Nat Commun 6:6756.

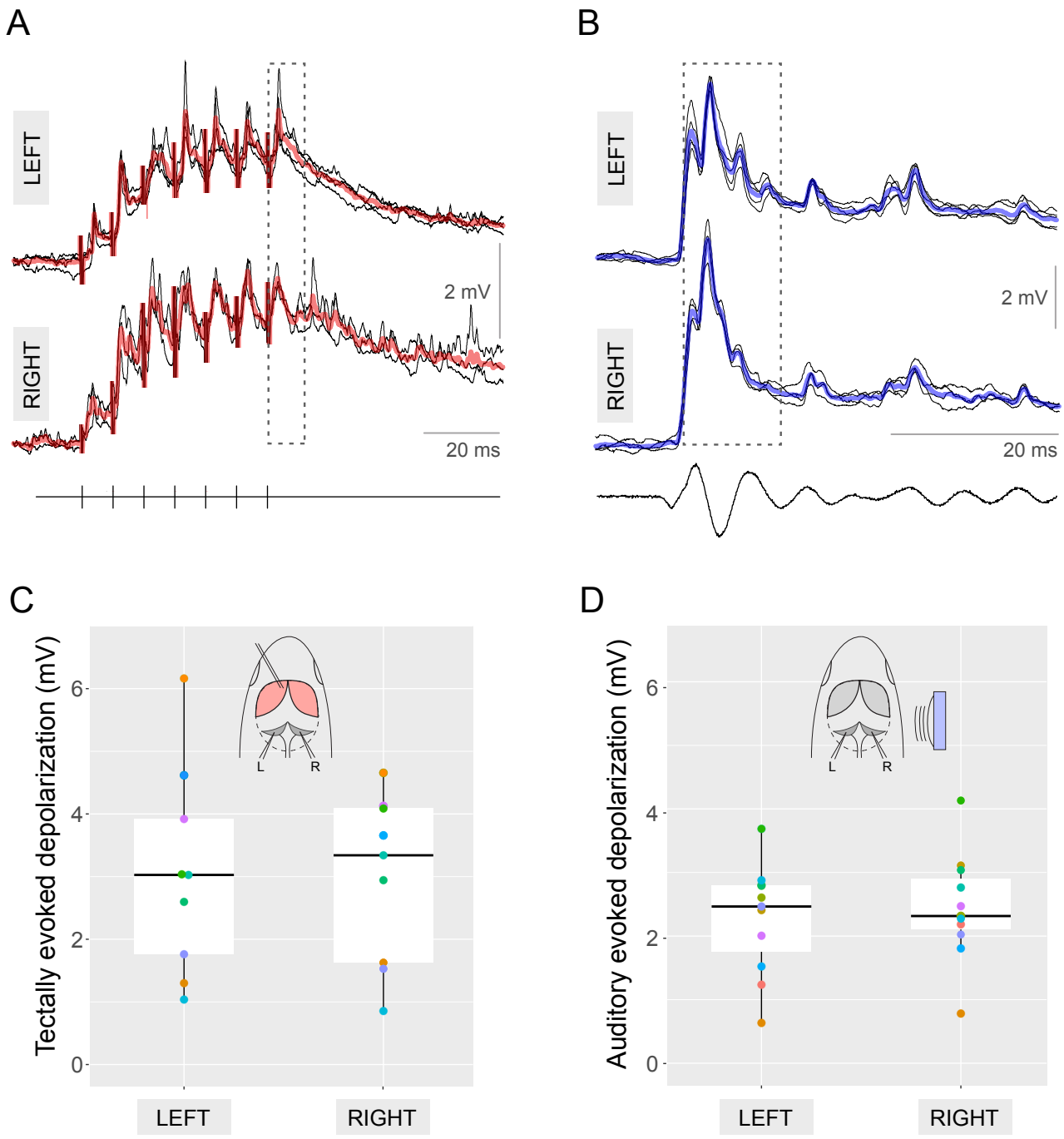
Weiss SA, Preuss T, Faber DS (2009) Phase encoding in the Mauthner system: implications in left-right sound source discrimination. J Neurosci 29:3431–3441.

Weiss SA, Zottoli SJ, Do SC, Faber DS, Preuss T (2006) Correlation of C-start behaviors with neural activity recorded from the hindbrain in free-swimming goldfish (Carassius auratus). J Exp Biol 209:4788–4801.

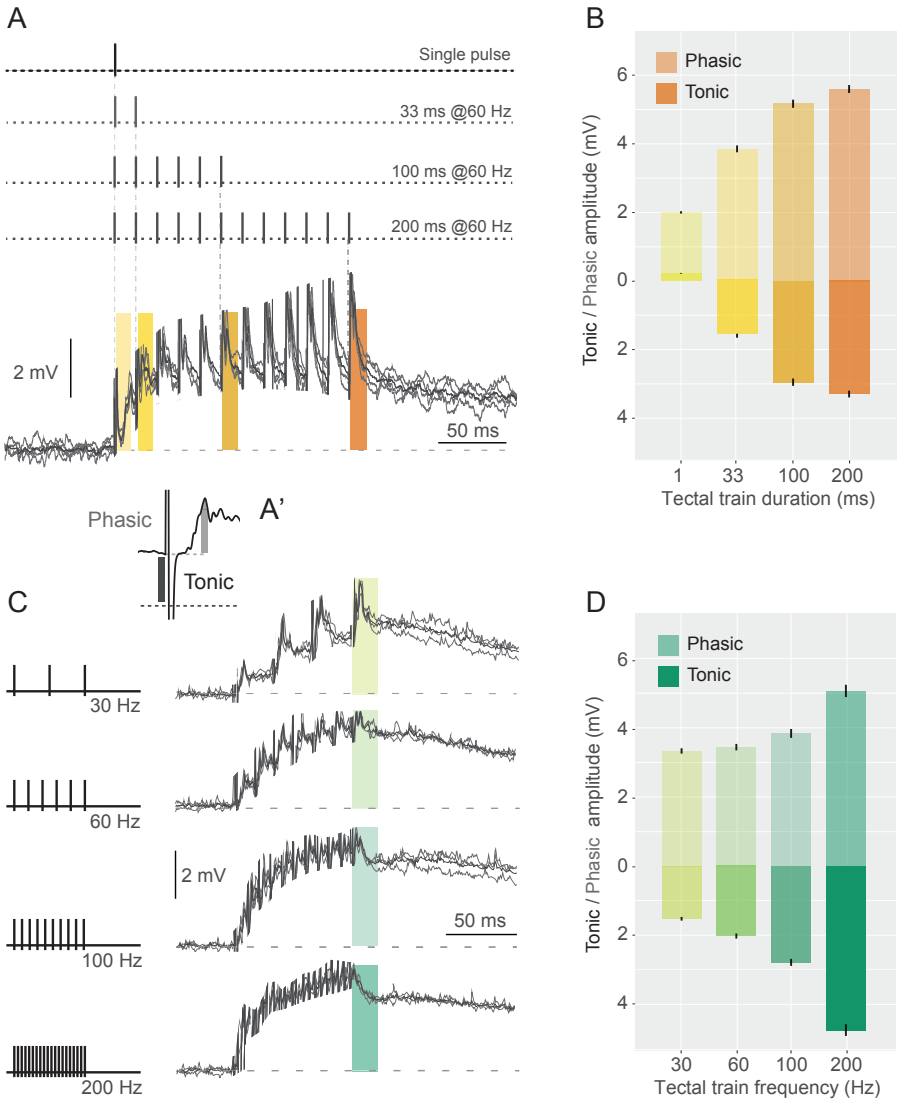
Zottoli SJ (1977) Correlation of the startle reflex and Mauthner cell auditory responses in unrestrained goldfish. J Exp Biol 66:243–254.

Zottoli SJ, Hordes AR, Faber DS (1987) Localization of optic tectal input to the ventral dendrite of the goldfish Mauthner cell. Brain Res 401:113–121.

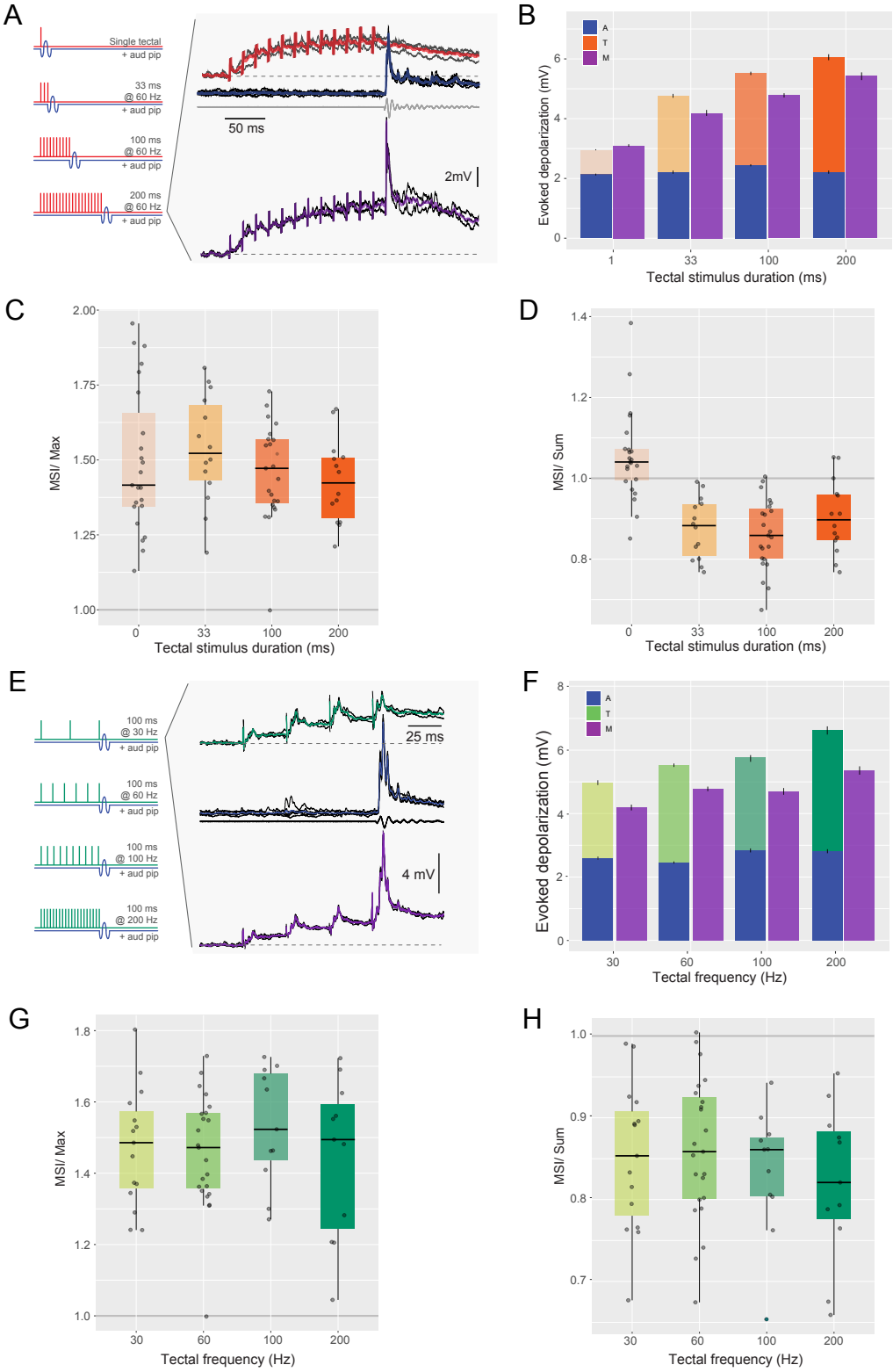
Supplementary Figure 1



# Figure 1


A
Single pulse
33 ms @60 Hz
100 ms @60 Hz
200 ms @60 Hz
2 mV
50 ms
Phasic
A'
Tonic
B
Phasic
Tonic
Tonic / Phasic amplitude (mV)
6
4
2
0
2
4
1
33
100
200
Tectal train duration (ms)
C
30 Hz
60 Hz
100 Hz
200 Hz
2 mV
50 ms
D
Phasic
Tonic
Tonic / Phasic amplitude (mV)
6
4
2
0
2
4
30
60
100
200
Tectal train frequency (Hz)

# Figure 2


A
Single tectal
+ aud pip
33 ms
@ 60 Hz
+ aud pip
100 ms
@ 60 Hz
+ aud pip
200 ms
@ 60 Hz
+ aud pip
50 ms
2mV
B
A
T
M
Evoked depolarization (mV)
Tectal stimulus duration (ms)
C
MSI/ Max
Tectal stimulus duration (ms)
D
MSI/ Sum
Tectal stimulus duration (ms)
E
100 ms
@ 30 Hz
+ aud pip
100 ms
@ 60 Hz
+ aud pip
100 ms
@ 100 Hz
+ aud pip
100 ms
@ 200 Hz
+ aud pip
25 ms
4 mV
F
A
T
M
Evoked depolarization (mV)
Tectal frequency (Hz)
G
MSI/ Max
Tectal frequency (Hz)
H
MSI/ Sum
Tectal frequency (Hz)

# Figure 3

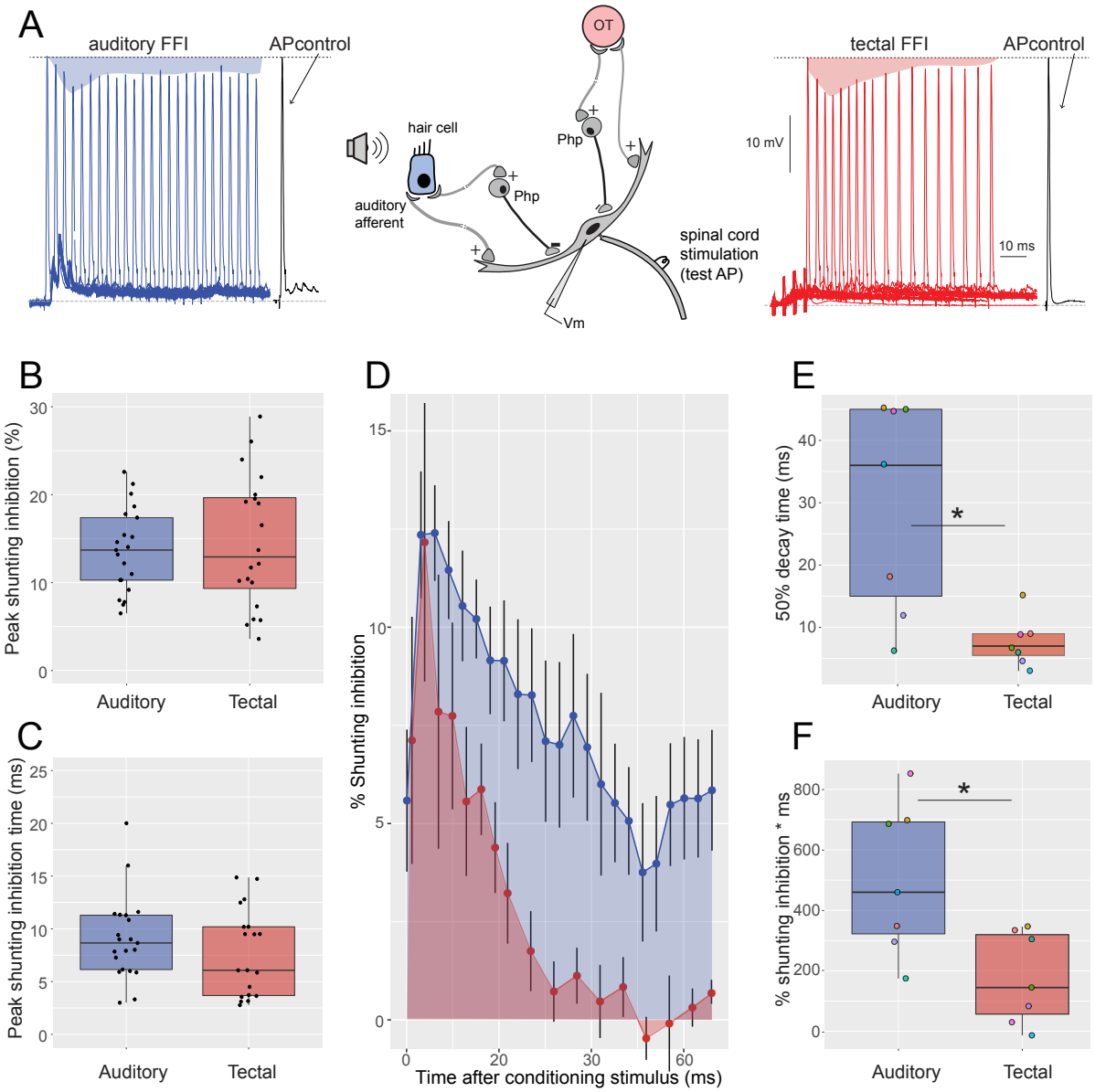

A
auditory FFI
APcontrol
hair cell
auditory afferent
Php
Php
OT
spinal cord stimulation (test AP)
Vm
tectal FFI
APcontrol
10 mV
10 ms
B
Peak shunting inhibition (%)
Auditory
Tectal
C
Peak shunting inhibition time (ms)
Auditory
Tectal
D
% Shunting inhibition
Time after conditioning stimulus (ms)
E
50% decay time (ms)
Auditory
Tectal
F
% shunting inhibition * ms
Auditory
Tectal

Figure 4

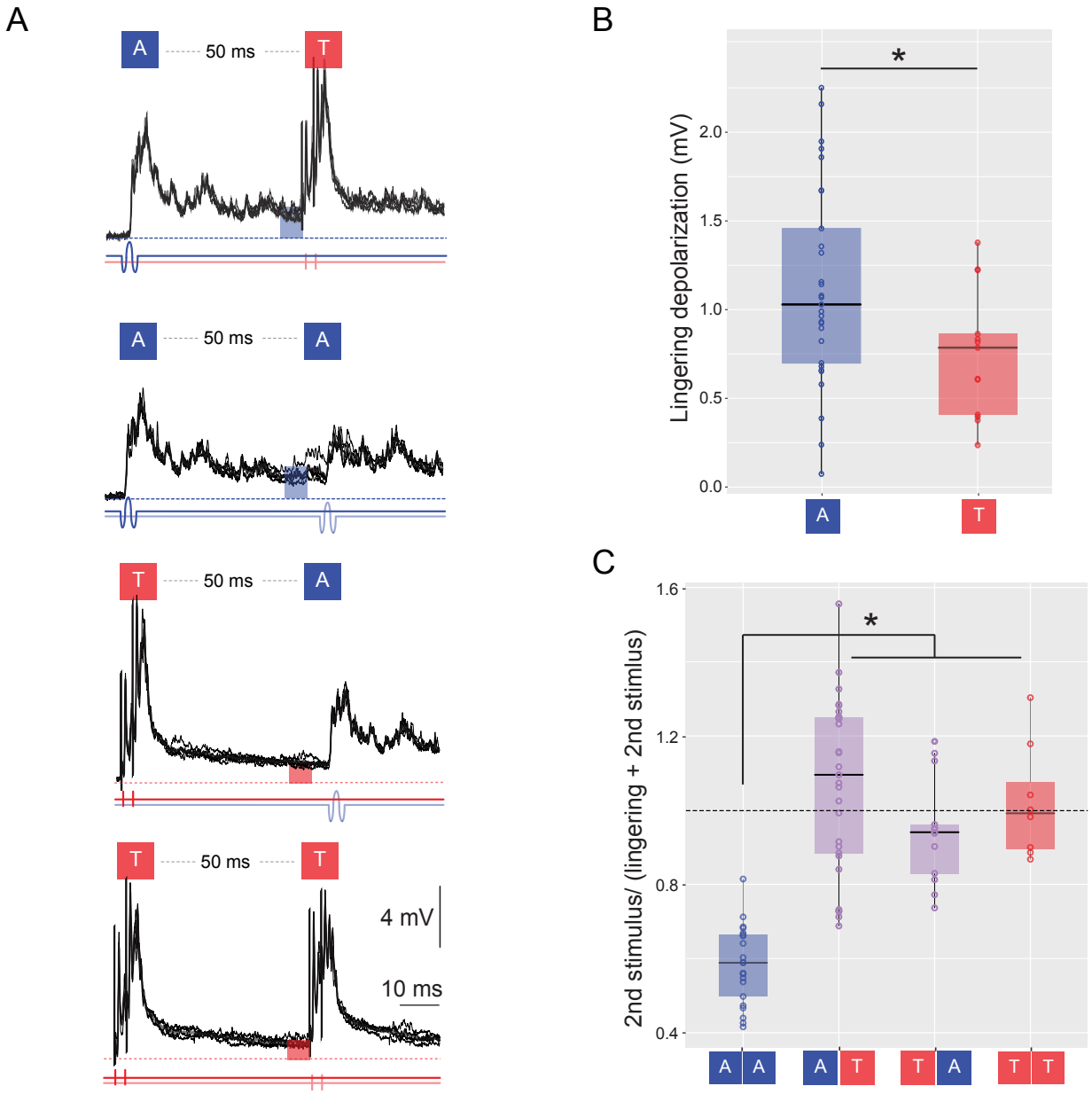
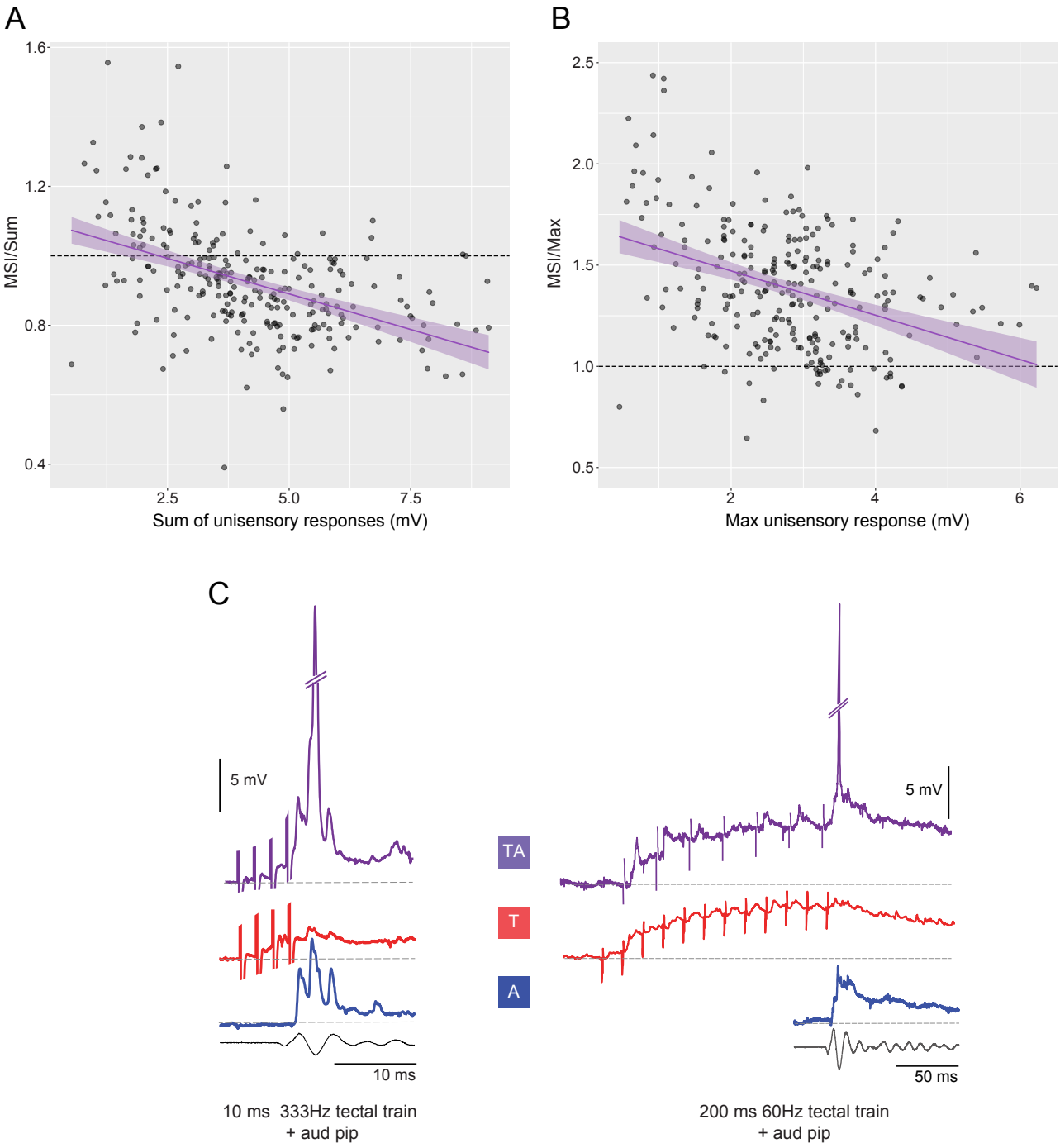

A
A
50 ms
T
A
50 ms
A
T
50 ms
A
T
50 ms
T
4 mV
10 ms
B
*
Lingering depolarization (mV)
2.0
1.5
1.0
0.5
0.0
A
T
C
*
2nd stimulus/ (lingering + 2nd stimlus)
1.6
1.2
0.8
0.4
A A
A T
T A
T T

Figure 5


A
1.6
1.2
0.8
0.4
MSI/Sum
2.5
5.0
7.5
Sum of unisensory responses (mV)
B
2.5
2.0
1.5
1.0
0.5
MSI/Max
2
4
6
Max unisensory response (mV)
C
5 mV
10 ms
TA
T
A
5 mV
50 ms
10 ms 333Hz tectal train
+ aud pip
200 ms 60Hz tectal train
+ aud pip