

1 **Perceptual salience is insufficient for auditory streaming**

2 **in eastern gray treefrogs (*Hyla versicolor*)**

3

4

5 Lata Kalra <sup>a,1</sup>, Shoshana Altman <sup>a</sup>, Mark A. Bee <sup>a</sup>

6 <sup>a</sup> Department of Ecology, Evolution, and Behavior, University of Minnesota, Saint Paul, MN, 55108, USA

7 <sup>1</sup>To whom the correspondence may be addressed: [kalra023@umn.edu](mailto:kalra023@umn.edu)

8 **Abstract**

9 Auditory streaming underlies a receiver's ability to organize complex mixtures of auditory input into  
10 distinct perceptual "streams" that represent different sound sources in the environment. During auditory  
11 streaming, sounds produced by the same source are integrated through time into a single, coherent  
12 auditory stream that is perceptually segregated from other concurrent sounds. Based on human  
13 psychoacoustic studies, a prominent hypothesis regarding auditory streaming is that *any* perceptually  
14 salient acoustic difference between sounds can promote their segregation into distinct auditory streams.  
15 Here, we used the eastern gray treefrog, *Hyla versicolor*, to test this hypothesis in the context of vocal  
16 communication in a non-human animal. In this system, females choose their mate based on perceiving  
17 features of a male's pulsatile advertisement calls in social environments (choruses) characterized by  
18 mixtures of overlapping vocalizations. We employed an experimental paradigm from human  
19 psychoacoustics to design interleaved pulsatile sequences (ABAB...) that mimicked key features of the  
20 species' advertisement call, and in which alternating pulses differed in pulse rise time, which is a robust  
21 species recognition cue in eastern gray treefrogs. Using phonotaxis assays, we found no evidence that  
22 perceptually salient differences in pulse rise time promoted the segregation of interleaved pulse  
23 sequences into distinct auditory streams. These results suggest the hypothesis that any perceptually  
24 salient acoustic difference can be exploited as a cue for stream segregation is not supported in all  
25 species. We discuss these findings in the context of cues used for species recognition and auditory  
26 streaming.

27 **Keywords**

28 Acoustic communication, Auditory perception, Auditory stream segregation, Gray treefrog, Species  
29 recognition, Vocal communication

## 30 Introduction

31 Acoustic communication, and hearing more generally, frequently requires listeners to perceive  
32 relevant sound sequences as distinct from other concurrent sounds (Cherry & Taylor, 1954; McDermott,  
33 2009). In humans, for example, following a conversation in noisy social settings (Remez, 2021; Repp,  
34 1988) or recognizing a melody in an orchestral piece (Dowling, 2012; McDermott & Oxenham, 2008)  
35 involves the ability to hear sound sequences (e.g., words, syllables, musical notes) as distinct from other  
36 sounds occurring at the same time. The ability to hear distinct sound sequences amid competing sounds  
37 is a non-trivial challenge because sounds from multiple sources sum to form a composite sound wave  
38 that impinges on the ears of a listener (Bregman, 1990). The composite sound wave must be perceptually  
39 organized into distinct “streams,” each corresponding to a coherent representation of the sound sequence  
40 produced by a given source. This process, called “auditory streaming” (Bregman, 1990), involves two  
41 complementary processes in which sounds produced by the same source are *integrated* into a coherent  
42 auditory stream while sounds produced by different sources are *segregated* into separate streams  
43 (Bregman, 1990; Bregman & Campbell, 1971; Moore & Gockel, 2012).

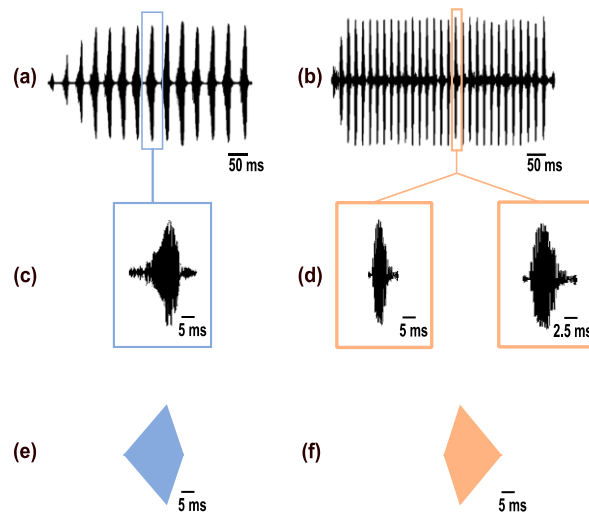
44 Psychoacoustic studies in humans have uncovered various cues influencing the integration  
45 versus segregation of sounds during auditory streaming. Many of these studies have employed a simple  
46 experimental paradigm wherein subjects listen to interleaved sequences of two types of tone pulses (A  
47 and B) and report their perception of the rhythm or rate of the sequence. The acoustic differences  
48 between the A and B pulses are manipulated across trials. Integration versus segregation can be  
49 assessed using this ABAB stimulus paradigm to determine whether subjects report hearing, as a function  
50 of the acoustic differences between the A and B pulses, a single, integrated sequence (ABAB...) or two  
51 segregated sequences (A–A–... and B–B–...), each at half the pulse rate of the actual stimulus sequence  
52 (van-Noorden, 1974). Sufficiently large differences in the spectral content (e.g., fundamental frequency or  
53 timbre), temporal patterns (e.g., onset/offset times, amplitude and frequency modulation patterns) or  
54 spatial location of A and B sequences promote their segregation, while smaller differences are more likely  
55 to result in their integration (reviewed in Bregman, 1990; Darwin, 1997, 2008; Micheyl & Oxenham, 2010;  
56 Shamma et al., 2011). The breadth of acoustic cues that facilitate auditory streaming in humans led  
57 Moore & Gockel (2012) to hypothesize that *any* perceptually salient difference between sound sequences  
58 can promote their segregation into separate streams.

59 Many non-human animals communicate using rhythmic sequences of sounds, such as pulsatile  
60 calls in frogs and crickets (Prestwich, 1994; Gerhardt & Huber, 2002), song motifs in songbirds and  
61 whales (Winn et al., 1981; Bruno & Tchernichovski, 2019), and echolocation clicks in bats and dolphins  
62 (Nihoul, 2004). Moreover, these signals are perceived in complex acoustic environments consisting of  
63 multiple biotic and abiotic sound sources (Bee & Micheyl, 2008; Gerhardt & Huber, 2002; Greenfield,  
64 2005). Auditory streaming is thus essential for accurate recognition, discrimination, and localization of  
65 signals across diverse species and behavioral contexts. Even though auditory streaming is a ubiquitous

66 communication challenge, the phenomenon has so far received little attention in studies of non-human  
67 animal communication (Hulse, 2002; Bee & Micheyl, 2008; Dent & Bee, 2018). Preliminary investigations  
68 using the ABAB paradigm in non-human animals suggest similar auditory streaming cues are used in  
69 humans and a diversity of other species. Frequency differences, for example, promote segregation in  
70 insects (J. Schul & Sheridan, 2006), frogs (Nityananda & Bee, 2011), fish (Fay, 2000; Fay, 1998), birds  
71 (MacDougall-Shackleton et al., 1998; Itatani & Klump, 2014; Dent et al., 2016), and mammals (Izumi,  
72 2002; Ma et al., 2010; Noda et al., 2013; Christison-Lagay & Cohen, 2014). Temporal differences in  
73 onset/offset times and amplitude modulation patterns promote segregation in frogs (Gupta & Bee, 2020)  
74 and birds (Itatani & Klump, 2009). Differences in spatial location promote segregation in insects (Dagmar  
75 von Helversen, 1984; Weber & Thorson, 1988), frogs (Farris et al., 2002; Farris et al., 2005; Bee, 2010)  
76 and mammals (Middlebrooks & Bremen, 2013; Yao et al., 2015). While these studies establish interesting  
77 parallels between auditory perception across taxa, it remains to be tested whether perceptual salience  
78 *per se* (sensu Moore & Gockel, 2012) is sufficient to promote segregation of sounds in non-human  
79 animals.

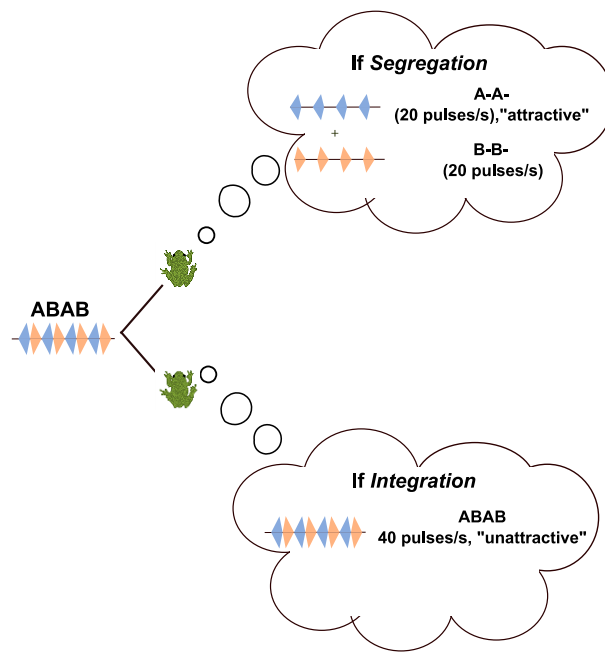
80           In this study of the eastern gray treefrog, *Hyla versicolor*, we used the ABAB stimulus paradigm  
81 to test the hypothesis that perceptually salient acoustic differences promote auditory streaming. The  
82 eastern gray treefrog is a well-studied frog in the context of animal communication that breeds in ponds  
83 and wetlands distributed throughout eastern North America (Gerhardt, 2001). Males of *H. versicolor*  
84 produce pulsatile advertisement calls (Fig. 1a) and breed in choruses. Even in small choruses of only  
85 conspecifics, there is a high degree of call overlap among neighboring males (Schwartz et al., 2002). In  
86 mixed-species choruses heterospecific males, including those of a morphologically indistinguishable  
87 sister species, *Hyla chrysoscelis*, also produce spectrally and temporally overlapping pulsatile  
88 advertisement calls (Fig. 1b) (Nityananda & Bee, 2011). Auditory streaming is thus crucial for female frogs  
89 to perceive the signal of a potential mate amidst other concurrent sounds (Bee, 2015). In *H. versicolor*,  
90 each advertisement call comprises of a sequence of 11-25 pulses (Fig. 1a). The amplitude time envelope  
91 of each pulse has a slow (approximately 65% of pulse duration) rise from pulse onset to peak amplitude  
92 and a fast (approximately 35% of pulse duration) fall from peak amplitude to pulse offset (Fig. 1c)  
93 (Gerhardt & Doherty, 1988; Ptacek et al., 1994; Gupta et al., 2021). Pulse amplitude rise and fall patterns  
94 – together described as “pulse shape” – facilitates species recognition in *H. versicolor*. Females of *H.*  
95 *versicolor* collected from a population in the State of Missouri in the central United States prefer pulses  
96 shaped with slow rise times, typical of conspecific calls (Fig. 1c), over pulses that have faster rise times  
97 and an overall shape that more closely resembles the heterospecific pulses of *H. chrysoscelis* (Fig. 1d).  
98 Rise time differences as small as 5 ms were perceptually salient and elicited strong behavioral  
99 discrimination between signals (Gerhardt & Schul, 1999). Here, we capitalized on this expected pulse rise  
100 time discrimination in *H. versicolor* to test the hypothesis that a perceptually salient difference in pulse  
101 rise time promotes the segregation of interleaved pulse sequences into separate auditory streams.

102 As a first step in our experimental design, we recorded and analyzed advertisement calls  
103 because pulse rise time preferences had not been established previously in our study population in the  
104 State of Minnesota. We used the results from our acoustic analyses to determine biologically realistic  
105 values of pulse rise time for subsequent behavioral experiments. We exploited a well-known preference  
106 for conspecific calls based on pulse-rate selectivity to design experiments based on the ABAB stimulus  
107 paradigm. The pulses in a *H. versicolor* call are, on average, about 20 ms long and separated by silent  
108 intervals of about 30 ms in duration. This regular rhythm corresponds to a pulse rate of 20 pulses/s  
109 (Gerhardt & Doherty, 1988; Gupta et al., 2021). Females of *H. versicolor* prefer the slower pulse rate of  
110 conspecific calls (Fig. 1a) over the faster pulse rate of *H. chrysoscelis* calls (Fig. 1b), which, on average,  
111 is approximately double (40 - 65 pulses/s) the pulse rate of conspecific calls (Noble & Hassler, 1936;  
112 Blair, 1958; Gerhardt, 1978; Ward et al., 2013). Accurate pulse-rate perception is crucial for species  
113 recognition as highlighted by the finding that two interleaved and identical conspecific pulse sequences  
114 are perceived as a single sequence with fast pulse rate that is less attractive to females of *H. versicolor*  
115 (Schwartz & Gerhardt, 1995; Schwartz & Marshall, 2006; see also Bee & Riemersma, 2008). Our use of  
116 the ABAB stimulus paradigm was based on a female's pulse-rate selectivity. We broadcasted two



**Fig. 1** Natural and synthetic signals of *H. versicolor* and *H. chrysoscelis*. **a** Oscillogram of a natural advertisement call of *H. versicolor*. **b** Oscillogram of a natural advertisement call of *H. chrysoscelis* depicting a faster pulse rate compared to *H. versicolor*. **c** A highlighted natural pulse of *H. versicolor* depicting a slow rise and relatively faster fall in amplitude. **d** left: A highlighted natural pulse of *H. chrysoscelis* (shown in the same time-scale as *H. versicolor* in panel c) depicting a fast rise and relatively slow fall in amplitude, and right: The same pulse magnified two-fold to highlight how the pulse shape (relative rise and fall patterns) in *H. chrysoscelis* is almost reversed relative to that of a natural *H. versicolor* pulse. **e** Synthetic “A” pulse (in blue) modelled on the overall duration and the rise and fall-times of a natural *H. versicolor* pulse. **f** Synthetic “B” pulse (in orange), which is a digitally reversed version of “A” pulse, has an overall duration typical of a natural *H. versicolor* and an overall shape typical of a natural *H. chrysoscelis* pulse.

117 interleaved sequences of pulses having the same conspecific pulse rate (each 20 pulses/s) but differing  
118 in their pulse rise time (A and B). The “A” pulses (Fig. 1e) had the pulse duration and rise time typical of  
119 conspecific pulses (Fig. 1c). The “B” pulses (Fig. 1f) were time-reversed versions of the “A” pulses, and  
120 therefore, had a pulse duration typical of conspecific pulses but an overall shape typical of heterospecific  
121 *H. chrysoscelis* pulses (Fig. 1d). We then measured stream segregation based on whether the subjects  
122 perceived two separate (A–A–... and B–B–...) sequences (indicating *segregation*), each with a preferred  
123 pulse rate (20 pulses/s), or a single (ABAB...) sequence (indicating *integration*) with a less preferred  
124 pulse rate (40 pulses/s) (Fig. 2). According to our hypothesis, we predicted subjects would be attracted to  
125 interleaved sequences that could be segregated into separate auditory streams, one of which was  
126 attractive, based on a perceptually salient difference in pulse rise time (Fig. 2).



**Fig. 2** Protocol for testing auditory streaming. We broadcasted interleaved pulsatile sequences of A (in blue) and B (in orange) pulses (ABAB) to female *H. versicolor*. If pulse rise time differences were sufficient to promote segregation of sounds, we expected the females to perceive two distinct sequences, A–A– and B–B–, both of which had the preferred conspecific pulse rate of 20 pulses/s and one of which (A–A–) also had the preferred conspecific pulse rise time. Consequently, the ABAB stimulus was predicted to be attractive to females. In contrast, if the rise time differences between the A and B pulses were insufficient to promote segregation, we would expect females to perceive a composite ABAB sequence as having a relatively less attractive pulse rate of 40 pulses/s.

127

## 128 **Methods**

### 129 **Subjects**

130 All acoustic recordings and behavioral tests were conducted using subjects from the Tamarack Nature  
131 Center (Ramsey County, MN, USA), which belong to the Midwest clade of *H. versicolor* (Booker et al.,  
132 2022). Acoustic recordings of males ( $n=30$ ) were made at night (between 2100 and 0100 h) in May and  
133 June of 2006 and 2021. For behavioral tests, females ( $n=43$ ) were collected in amplexus at night  
134 (between 2100 and 0100 h) in May and June of 2021. Amplexed pairs were returned to the laboratory  
135 where they were maintained at approximately 4°C to delay egg laying and maintain behavioral  
136 responsiveness (Gerhardt, 1995). Prior to behavioral testing, frogs were placed in an incubator for at least  
137 30 minutes and allowed to reach a body temperature of 20°C. Between trials, females were returned to  
138 the 20°C incubator with their mates for a minimum of 5 min to maintain body temperature and preserve  
139 responsiveness. Because *H. versicolor* breeds syntopically with *H. chrysoscelis* at our field site, we  
140 confirmed the species identity of all subjects in an initial two-alternative choice test in which we  
141 broadcasted alternating synthetic models of the two species' calls (as in Gupta et al., 2021). Only females  
142 that approached the *H. versicolor* stimulus were used as subjects in the experiments described below. In  
143 some case, females were also used as subjects for other experiments not described here. There is little  
144 evidence for "carryover" effects between consecutive phonotaxis tests separated by several minutes  
145 (Akre & Ryan, 2010; Gerhardt, 1981). All frogs were released at their collection site within 48 hours of  
146 completing behavioral tests.

### 147 **Acoustic recordings and analysis**

148 Vocalizations were recorded (44.1 kHz sampling rate, 16-bit resolution) using Sennheiser ME66 or ME67  
149 microphones (Sennheiser USA, Old Lyme, CT, U.S.A.) connected to Marantz PMD620 or PMD670  
150 recorders (D&M Professional, Itasca, IL, U.S.A.). Microphones were held by hand or mounted on a tripod,  
151 and the tip of the microphone was positioned approximately 1 m away from the focal male. For each  
152 individual male ( $n = 30$ ) we recorded and analyzed a minimum of 5 calls (range, 5 - 45 calls/male). Since  
153 both the acoustic properties of advertisement calls and female preferences for call properties can vary  
154 with temperature (Gerhardt, 1978), we measured the wet-bulb air temperature and the water temperature  
155 at each male's calling site immediately following each recording. We noted the general position from  
156 which the male was calling (e.g., in air on emergent vegetation versus floating on the surface of the  
157 water) to determine the most appropriate temperature for later use to standardize call properties to a  
158 common temperature of 20°C. We recorded males from two different ponds and from different areas  
159 within each pond across nights to reduce the chances of recording the same individual multiple times.

160 Acoustic recordings were analyzed using SoundRuler version 0.9.6.0 (Gridi-Papp, 2007), which  
161 performs automatic recognition of small repeated acoustic elements and exports an output summary of  
162 numerous acoustic properties (Bee, 2004). The output summary was further analyzed in R studio (R Core

163 Team, 2020) to derive and analyze specific acoustic properties of interest for which we computed means,  
164 standard deviations (SD), and ranges. Our primary focus was on pulse shape, which we characterized by  
165 measuring pulse rise time (ms, time from pulse onset to peak amplitude) and pulse fall time (ms, time  
166 from pulse peak amplitude to offset). To place measures of pulse shape in the overall context of the  
167 advertisement call, we also measured other temporal properties, including pulse duration (ms), pulse rate  
168 (pulses/s), call duration (pulses/call), call rate (calls/min), and spectral properties including the frequency  
169 (Hz) of each pulse's first and second harmonics, which correspond to the fundamental frequency and  
170 dominant frequency, respectively. Because the recordings were made at different temperatures, we  
171 followed Platz & Forester (1988) to standardize all call properties to 20°C, which is close to the average  
172 temperature observed in our recordings as well as the temperature at which we performed behavioral  
173 experiments.

#### 174 **Acoustic stimuli**

175 Synthetic acoustic stimuli (44.1 kHz sampling rate, 16-bit resolution) were generated in MATLAB R2020b  
176 (Mathworks, Natick, MA, USA) using parameter values taken from our acoustic analysis of natural  
177 signals. Across all experiments, stimuli were designed to stimulate a calling male and were constructed  
178 as a 5 min sequence of synthetic calls that repeated at a rate of 10 calls/min. Each call was generated as  
179 a sequence of pulses wherein each pulse was 20 ms long and composed of two phase-locked spectral  
180 components (1250 Hz and 2500 Hz, corresponding to the fundamental and dominant frequencies,  
181 respectively, of the natural signals). Further, based on our acoustic analysis, the amplitude of the 1250 Hz  
182 component was fixed to be 13 dB lower relative to the 2500 Hz component. The calls within each stimulus  
183 sequence differed in the rise time, rate, and timing of their constituent pulses according to the type of  
184 phonotaxis test performed, as described next.

#### 185 **Experimental design**

186 We performed three different choice tests, described below, using female phonotaxis as a behavioral  
187 assay. Each test was replicated twice using stimuli presented at one of two different sound pressure  
188 levels (80 dB peak SPL and 100 dB peak SPL). We used these sound pressure levels because auditory  
189 perception in frogs can be sound-level dependent (Gerhardt, 1987, 2005, 2008) and because these levels  
190 encompass much of the natural range of variation in the sound pressure levels of advertisement calls  
191 (Gerhardt, 1975). Each subject was tested in each of the six tests (3 choice tests × 2 signal levels) in a  
192 randomized order. We used two-tailed binomial tests to compare the proportion of females choosing a  
193 specific stimulus to the chance expectation if they chose randomly. All data analysis was performed in R  
194 studio (R Core Team, 2020).

195 **Perceptual salience test.** We performed a two-alternative choice test to determine whether pulse rise  
196 time differences are perceptually salient in our study population. This test simulated a choice between two

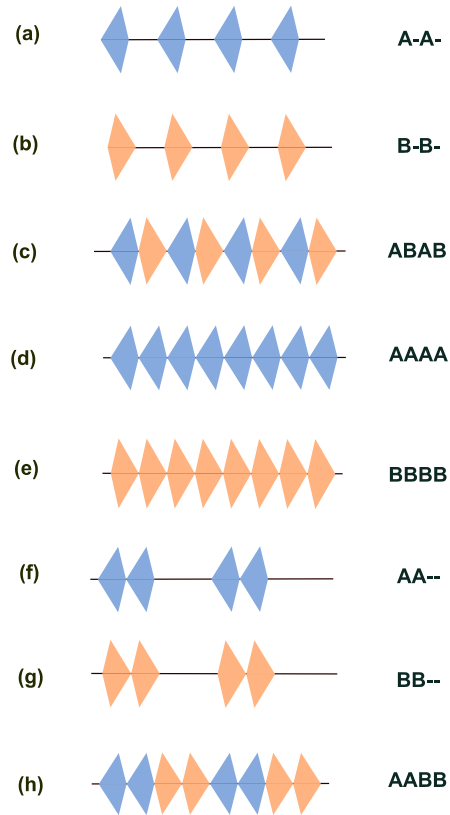


197 calling males producing calls having attractive pulse rates of 20 pulses/s, with each call comprising 16  
198 pulses that differed only in pulse rise time (“A” versus “B”). “A” pulses (Fig. 1e) had slow, linear rise times  
199 (13 ms, 65% of pulse duration) and fast, linear fall times (7 ms, 35% of pulse duration). The “A–A–”  
200 stimulus (Fig. 3a) consisted of pulses that simulated the average rise time of conspecific pulses in *H.*  
201 *versicolor* as determined in our acoustic analysis (Noble & Hassler, 1936; Blair, 1958; Gerhardt, 1978). In  
202 contrast, the “B–B–” stimulus (Fig. 3b) consisted of time-reversed A pulses that had fast, linear rise times  
203 (7 ms, 35% of pulse duration) and slow, linear fall times (13 ms, 65% of pulse duration). These “B” pulses  
204 (Fig. 1f) closely resembled the overall shape of pulses in the heterospecific calls of *H. chrysoscelis* (fast  
205 pulse rise ~ 35% of pulse duration and slow pulse fall ~58% of pulse duration; Ward et al., 2013). By  
206 using time-reversed “A” pulses in the “B–B–” stimulus, we ensured both stimuli had pulses of consistent  
207 duration and peak sound pressure levels and differed only in pulse rise time and fall time (Diekamp &  
208 Gerhardt, 1995). Based on expectations from previous work by Gerhardt and Schul (1999) in Missouri  
209 populations, we predicted a proportion of subjects significantly higher than 0.5 would choose the A–A–  
210 stimulus (Fig. 4a). As discussed below, this prediction was supported by the data, thus allowing us to use  
211 the perceptually salient differences between the “A” and “B” pulses to test an auditory streaming  
212 hypothesis.

213 **Auditory streaming test.** We used a four-alternative choice test to evaluate the hypothesis that females  
214 of *H. versicolor* can use perceptually salient difference in pulse rise time to segregate temporally  
215 overlapping calls into separate auditory streams. The key stimulus was based on the ABAB stimulus  
216 paradigm. It was created by temporally interleaving pulses from the A–A– and B–B– stimuli from the  
217 perceptual salience test to produce an “ABAB” stimulus (Fig. 3c). This ABAB stimulus had 32 pulses and  
218 a composite pulse rate of 40 pulses/s (simulating a less attractive *H. chrysoscelis* call) but was made up  
219 of two component pulse sequences (A–A– and B–B–) each having an attractive conspecific pulse rate of  
220 20 pulses/s, only one of which (A–A–) also had the more attractive pulse rise time of conspecific calls.  
221 Two other stimuli in this four-alternative choice test also had the less preferred pulse rate of 40 pulses/s  
222 and consisted of a sequence of either all A pulses (“AAAA,” Fig. 3d) or all B pulses (“BBBB,” Fig. 3e). The  
223 final stimulus was created by interleaving pairs of A (“AA—,” Fig. 3f) and B (“BB—,” Fig. 3g) pulses so  
224 that, like the ABAB stimulus, this “AABB” (Fig. 3h) stimulus consisted of two pulse sequences having  
225 average pulse rates of 20 pulses/s. The main difference between the ABAB and AABB stimuli was that  
226 the former was comprised of two interleaved sequences having “regular” pulse timing (A–A– and B–B–),  
227 as determined by their constant 30 ms inter-pulse interval between consecutive pulses (typical of natural  
228 advertisement calls), whereas the component sequences in the AABB stimulus had “irregular” pulse  
229 timing (AA— and BB—) created by having inter-pulse intervals that alternated between 10 ms and 50 ms  
230 between consecutive pulses but averaged to 30 ms over the duration of each composite stimulus. Among  
231 all the pulse sequences used across the four stimuli, the A–A– component sequence in the ABAB  
232 stimulus was expected to be the most attractive because it was the only stimulus with the pulse rise  
233 times, pulse rate, and pulse timing typical of conspecific calls (Gerhardt & Doherty, 1988; Gerhardt &

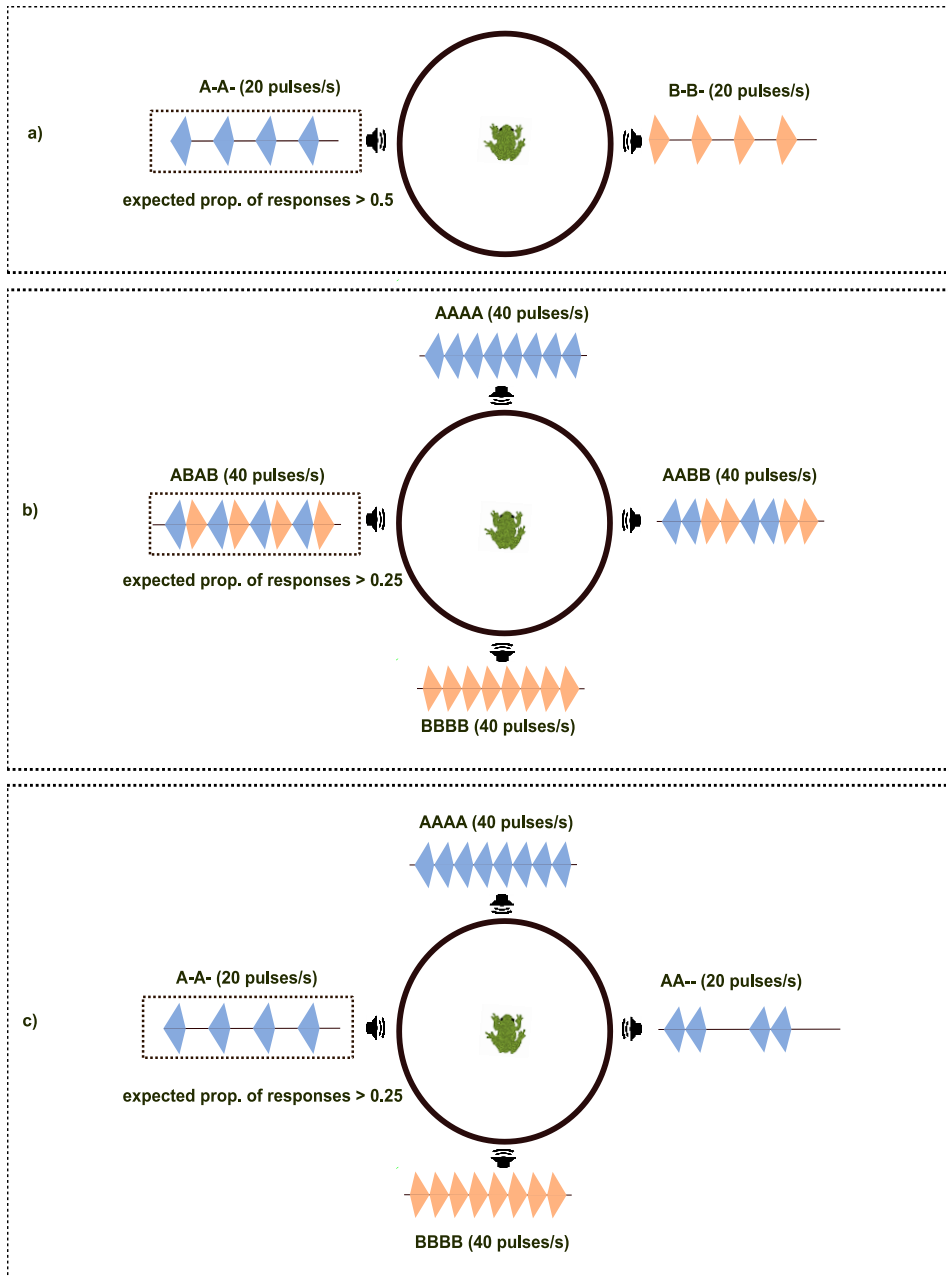
234 Schul, 1999; Gerhardt, 2005). If the perceptually salient difference between the A and B pulse rise times  
235 was sufficient to allow auditory streaming, we predicted females would be attracted to the A–A–  
236 component of the ABAB stimulus and thus choose the ABAB stimulus over the other three stimuli, which  
237 had less preferred pulse rates (AAAA, BBBB, and AABB), pulse rise times (BBBB and AABB), or pulse  
238 timing (AABB). Therefore, if auditory streaming of the interleaved A and B pulses in the ABAB stimulus  
239 occurred, we predicted that the proportion of subjects choosing the ABAB stimulus would be significantly  
240 higher than 0.25 (Fig. 4b).

241 **Pulse-rate and pulse-timing test.** We performed a final four-alternative choice test to confirm that  
242 females in our population were selective for conspecific pulse rates and regular pulse timing. The A–A–  
243 stimulus had a conspecific pulse rate of 20 pulses/s, the conspecific pulse rise time (A), and regular pulse  
244 timing. The AAAA and BBBB stimuli both had a faster pulse rate of 40 pulses/s (typical of the  
245 heterospecific calls of *H. chrysoscelis*) and regular pulse timing but differed in having either conspecific  
246 (A) or heterospecific (B) pulse rise times. Finally, the AA— stimulus had a conspecific pulse rate of 20  
247 pulses/s (on average), the conspecific pulse rise time (A), but irregular pulse timing (alternating 10 ms  
248 and 50 ms inter-pulse intervals). We predicted that if females prefer calls with conspecific pulse rates and  
249 evenly spaced pulses – two key provisions of our test of auditory streaming – then they would choose the  
250 A–A– stimulus at a rate significantly higher than the chance proportion of 0.25 (Fig. 4c).



**Fig. 3** Acoustic stimuli for behavioral experiments. **a** “A-A-,” and **b** “B-B-” were constructed as sequences of A and B pulses, respectively, repeating at a rate of 20 pulses/s and having a regular inter-pulse interval of 30 ms. **c** “ABAB” was constructed by temporally interleaving the “A-A” and “B-B-” sequences and had a composite rate of 40 pulses/s. **d** “AAAA,” and **e** “BBBB” were constructed as sequences of A and B pulses, respectively, repeating at a rate of 40 pulses/s. **f** “AA—,” and **g** “BB—,” were constructed as sequences of A and B pulses, respectively, repeating at an average rate of 20 pulses/s and had an irregular inter-pulse interval that shifted between 10 ms and 50 ms between consecutive pulses. **f** “AABB” was constructed by temporally interleaving the “AA—” and “BB—” sequences and had a composite rate of 40 pulses/s.

252



**Fig. 4** Design and predictions for the behavioral experiments. **a** Perceptual salience test. If the rise time differences between pulses A and B are perceptually salient, subjects were expected to prefer A-A- stimulus more than the chance probability of 0.5. **b** Auditory streaming test. If perceptually salient rise time differences are sufficient to allow auditory streaming, subjects were expected to prefer ABAB stimulus more than the chance probability of 0.25. **c** Pulse-rate and pulse-timing test. If subjects prefer calls with conspecific pulse rates and evenly spaced pulses, they were expected to prefer A-A- stimulus more than the chance probability of 0.25.

253

## 254 **Testing protocol**

255 Behavioral tests were performed in a 2-m diameter circular phonotaxis arena surrounded by a 60-cm tall  
256 wall. The arena wall was constructed from hardware cloth and black fabric to create a visually opaque but  
257 acoustically transparent barrier. The arena was set within a hemi-anechoic sound chamber (length ×  
258 width × height: 2.8 × 2.3 × 2.1 m; Industrial Acoustics Company, IAC, North Aurora, IL, USA). Stimuli  
259 were broadcast from an HP ProBook 450 G6 (HP inc., Palo Alto, CA, USA) through a MOTU M4 sound  
260 card (MOTU, Inc., Cambridge, MA, USA) using Adobe Audition 3.0 (Adobe Systems Inc. San Jose, CA,  
261 USA). The output audio was amplified by a Crown XLS 1000 High-Density Power Amplifier (Crown  
262 International, Los Angeles, CA, USA) and played through one of four Orb1 speakers (Orb Audio,  
263 Sherman Oaks, CA, USA) located outside the arena wall on the floor of the sound chamber. The four  
264 speakers were evenly spaced and separated by 90° around the circumference of the circular test arena  
265 and positioned to face inward toward the center of the arena. The sound pressure level (SPL, LCpeak, re  
266 20 µPa) of stimuli broadcast through each speaker was measured for calibration using a sound level  
267 meter (Larson Davis Model 831, Larson Davis Inc., Depew, NY) attached to a microphone placed at the  
268 center of the arena at the same level above the floor as a subject's ears and aimed toward the speaker.  
269 For four-choice tests, alternative stimuli were broadcasted through four different speakers while for two-  
270 choice tests, alternative stimuli were broadcasted through two speakers located 180° apart. The  
271 consecutive calls within each stimulus sequence were separated by a silence of 5.23 s. The timing of  
272 calls across all stimulus sequences was such that an equal period of silence preceded and followed each  
273 call to avoid call overlap and any leader/follower relationships among stimulus calls. We further controlled  
274 for any leader/follower relationships by randomizing, across subjects, the order in which the very first calls  
275 of different stimuli were broadcast in a playback.

276 At the start of each test a single subject was placed at the center of the circular arena inside an  
277 acoustically transparent release cage. After a 60-s acclimation period, stimulus broadcasts began. After  
278 two repetitions of each test stimulus, a timer was started, and the frog was remotely released by lifting the  
279 lid of the release cage using a pulley system that could be operated from outside the sound chamber.  
280 Subjects' responses were monitored through an overhead IR camera mounted directly over the test  
281 arena. Subjects were given up to 5 min to respond. A response was recorded if a subject approached to  
282 within 10 cm of a speaker and remained there for 30 s. A no-response was recorded if a frog failed to exit  
283 the release cage within 3 min after its release or if it failed to meet our response criterion within 5 min.

## 284 **Results**

285 **Call analyses.** The mean ( $\pm$  SD) rise and fall times of *H. versicolor* pulses were 13.0 ms ( $\pm$  2.6 ms; range  
286 = 7.9 - 19.6 ms) and 7.4 ms ( $\pm$  1.8 ms; range = 4.4 - 14.1 ms), respectively. The mean pulse duration was  
287 20.4 ms ( $\pm$  3.1 ms; range = 13.8 - 26.8 ms). Thus, on average, the pulse rise and fall times, respectively,  
288 were close to 65% and 35% of the call duration. Descriptive statistics for all other acoustic properties are

289 reported in Table 1. Based on these results, we chose the rise and fall times of “A” pulses as 13 ms and 7  
290 ms respectively. Since “B” pulses were digitally reversed versions of “A” pulses, the rise and fall times of  
291 “B” pulses were 7 ms and 13 ms, respectively.

292 **Perceptual salience test.** In the two-alternative choice test comprising the perceptual salience test,  
293 approximately, 98% (42 out of 43) and 91% (39 out of 43) of the females tested responded by making a  
294 choice at sound pressure levels of 100 dB and 80 dB, respectively. We predicted that females would  
295 prefer signals having a conspecific pulse rate and slow pulse rise time (A–A–) over an alternative having  
296 a conspecific pulse rate but a fast pulse rise time, and overall shape typical of heterospecific *H.*  
297 *chrysoscelis* pulse (B–B–). The data were consistent with this prediction. The proportion of subjects  
298 choosing A–A– stimulus over the B–B– stimulus was significantly higher than expected by chance (0.50)  
299 at both signal levels. At 100 dB, approximately 95% of subjects chose A–A– ( $n = 42$ ,  $p < 0.001$ ) and at 80  
300 dB, approximately 92% of subjects chose A–A– ( $n = 39$ ,  $p < 0.001$ ) (Fig. 5a). The observed behavioral  
301 discrimination based on pulse rise time confirmed that differences in pulse rise time were both  
302 perceptually and behaviorally salient.

303 **Auditory streaming test.** In the four-alternative choice test (ABAB vs. AAAA vs. BBBB vs. AABB) to  
304 investigate auditory streaming, approximately 77% (33 out of 43) and 70% (30 out of 43) of the females  
305 tested responded by making a choice at sound pressure levels of 100 dB and 80 dB, respectively. We  
306 predicted females would be attracted to the A–A– component of the ABAB stimulus if auditory streaming  
307 based on pulse rise time differences occurred. The data were not consistent with this prediction. The  
308 proportion of subjects choosing ABAB was not significantly higher than the chance probability of 0.25  
309 (exact binomial test,  $\alpha = 0.05$ ) at either of the signal levels. At 100 dB, approximately 15% of subjects  
310 chose ABAB ( $n = 33$ ,  $p = 0.231$ ) and at 80 dB, approximately 37% of subjects chose ABAB ( $n = 30$ ,  $p =$   
311 0.143) (Fig. 5b).

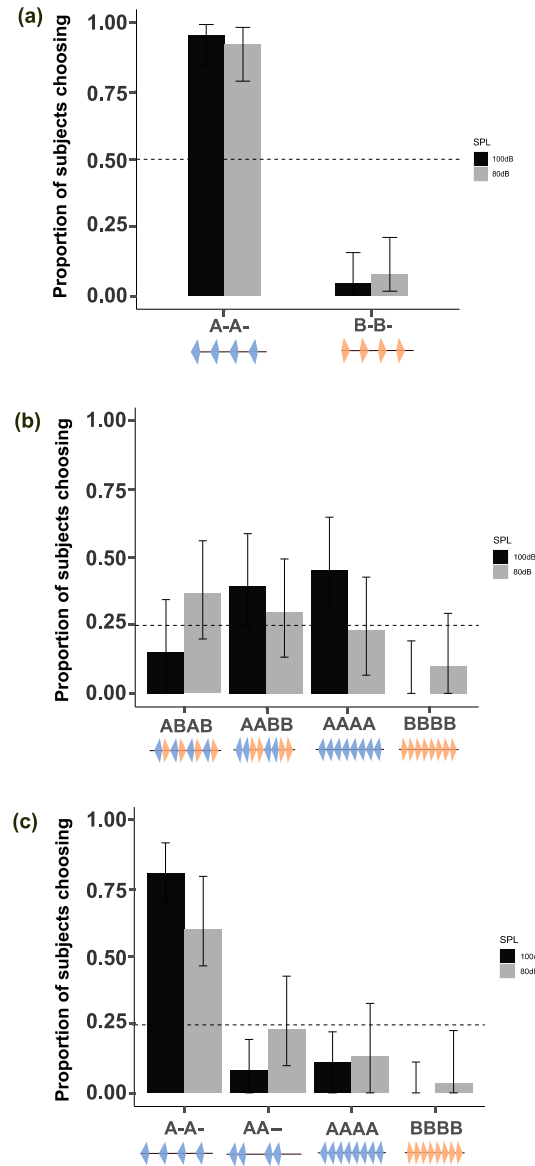
312 **Pulse-rate and pulse-timing test.** In the four-alternative choice test (A–A– vs. AAAA vs. BBBB vs. AA—  
313 ) to confirm pulse-rate and timing preferences, approximately 84% (36 out of 43) and 72% (31 out of 43)  
314 of the females tested responded by making a choice at sound pressure levels of 100 dB and 80 dB,  
315 respectively. We predicted that females would prefer the stimulus with a conspecific pulse rate and  
316 regular pulse timing (A–A–) over those with heterospecific pulse rates (AAAA and BBBB) and irregular  
317 pulse timing (AA—). The data were consistent with this prediction. The proportion of subjects choosing  
318 A–A– (slow rate and regular timing) was significantly higher than a chance probability of 0.25 (exact  
319 binomial test,  $\alpha = 0.05$ ) at both signal levels. The percentage of subjects choosing A–A– was  
320 approximately 81% at 100 dB ( $n = 36$ ,  $p < 0.001$ ) and 61% at 80 dB ( $n = 31$ ,  $p < 0.001$ ) (Fig. 5c).

321

322

323 **Table 1:** Descriptive statistics of acoustic properties of *H. versicolor* advertisement calls ( $n = 30$  males)  
324 recorded in Minnesota and standardized to a temperature of 20° C. The range of temperatures at which  
325 males were recorded was 10.2 – 29.0° C.

<b>Acoustic property</b>	<b>mean <math>\pm</math> SD</b>	<b>Range</b>
Pulse rise time (ms)	13.0 $\pm$ 2.6	7.9 - 19.6
Pulse fall time (ms)	7.4 $\pm$ 1.8	4.4 - 14.1
Pulse rate (pulses/s)	19.3 $\pm$ 3.1	14.9 - 25.3
Pulse duration (ms)	20.4 $\pm$ 3.1	13.8 - 26.8
Call duration (pulses/call)	15.9 $\pm$ 3.3	11.1 – 25.1
Call rate (calls/min)	14.1 $\pm$ 4.0	5.1 – 22.2
Pulse fundamental frequency (Hz)	1232.5 $\pm$ 85.5	1084.7 – 1501.2
Pulse dominant frequency (Hz)	2465.1 $\pm$ 170.9	2169.5 – 3002.4



**Fig. 5** Results for behavioral experiments. Black and gray bars indicate the proportions of subjects choosing a given stimulus at 100 dB and 80 dB, respectively. **a** Results for the perceptual salience test. Error bars depict exact 95% binomial confidence intervals (CIs). **b** Results for the auditory streaming experiment. Error bars depict 95% multinomial CIs. **c** Results for the pulse-rate and pulse-timing test. Error bars depict 95% multinomial CIs. Horizontal dashed lines depict the chance probability for each experiment.

326

327



## 328 Discussion

329 The goal of this study was to test the hypothesis (sensu Moore & Gockel, 2012) that perceptual salience  
330 *per se* is sufficient to promote auditory streaming in non-human animals. Our results are inconsistent with  
331 this hypothesis. A species-typical difference in pulse rise time was perceptually salient, as evidenced by  
332 strong behavioral discrimination based on this acoustic cue in two-choice tests. However, there was no  
333 evidence that this salient acoustic cue also promoted the perceptual segregation of two interleaved  
334 pulses sequences differing in pulse rise time. Based on this outcome, we provisionally conclude that  
335 perceptual salience was insufficient for auditory streaming in the context of segregating temporal  
336 sequences of pulses in overlapping calls in *H. versicolor*.

337 Our bioacoustic analyses confirmed the presence of species differences in pulse rise time  
338 between *H. versicolor* and its sister species, *H. chrysoscelis*, in Minnesota that were similar to differences  
339 reported in other populations (Gerhardt & Doherty, 1988). Pulses in *H. versicolor* had rise times (range:  
340 7.9 to 19.6 ms; Table 1) that were, on average, about 10 ms slower than those in the calls of *H.*  
341 *chrysoscelis* recorded in the same geographic area (range: 1.8 to 4.7 ms; Ward et al., 2013). Moreover,  
342 our two-choice test of perceptual salience demonstrated that a rise time difference of just 6 ms was  
343 perceptually salient and elicited a robust preference (by 92 – 95% subjects) for slow rise times. This  
344 finding corroborates previous work on pulse rise time preferences in female *H. versicolor* from a Missouri  
345 population (Gerhardt & Doherty, 1988; Gerhardt & Schul, 1999). Both the absolute rise times between  
346 the A and B pulses used in our study (13 ms versus 7 ms, respectively), and their relative difference (6  
347 ms) were close to those tested by Gerhardt & Schul (1999; e.g., 12.5 ms versus 7.5 ms). Our study used  
348 20-ms pulses with two spectral components (1200 Hz and 2400 Hz), whereas Gerhardt & Schul (1999)  
349 used 25-ms pulses having just the lower or the higher spectral component alone. In both studies, females  
350 of *H. versicolor* rejected a fast rise time more typical of the pulses in calls produced by male *H.*  
351 *chrysoscelis* in favor of a slow rise time typical of the pulses in conspecific calls. Our findings add to the  
352 evidence that pulse rise time, along with other fine temporal features like pulse rate, facilitate pre-mating  
353 species isolation between *H. versicolor* and *H. chrysoscelis*, which have spectrally similar calls (Gerhardt,  
354 2005). As such, the present findings also contribute to our current understanding of how signal  
355 preferences may persist or vary across different populations and geographical lineages of treefrogs (e.g.,  
356 Gerhardt et al., 2007; Gupta & Bee, 2023).

357 Despite strong behavioral discrimination between two pulse sequences differing in pulse rise time  
358 (i.e. A–A– versus B–B–), there was no evidence that the same difference promoted auditory streaming  
359 when the same sequences were temporally interleaved (i.e., ABAB). Females did not prefer the ABAB  
360 stimulus when it was presented in a four-choice test with alternatives having less preferred pulse rates,  
361 pulse rise times, or pulse timing (AAAA, BBBB, and AABB). This result is contrary to our prediction that  
362 the rise time difference would promote segregation of the ABAB sequence into separate streams, one of  
363 which corresponded to a pulse sequence (A–A–) having the preferred pulse rate, pulse rise time, and

364 pulse timing typical of conspecific calls. One alternative explanation for this lack of preference could be  
365 that subjects perceptually segregated the ABAB stimulus into separate A–A– and B–B– streams based  
366 on rise time differences but behaviorally avoided the source of the perceived B–B– stream consisting of  
367 pulses with fast rise times. This explanation seems unlikely for several reasons based on other work in  
368 this species. First, females of *H. versicolor* will approach the calls of a male *H. chrysoscelis* in a no-choice  
369 test when it is the only stimulus presented, suggesting stimuli with both fast pulse rates and fast pulse rise  
370 times are not inherently aversive (Gerhardt & Doherty, 1988). Second, Bush et al. (2002) and Schul &  
371 Bush (2002) showed that females responded in no-choice tests to a broad range of stimuli having  
372 different pulse rise times, including rise times faster than those of the B pulses in our stimuli. Third,  
373 Gerhardt et al. (1994) showed that females of *H. versicolor* did not avoid *H. chrysoscelis* calls while  
374 approaching a conspecific call, and Schwartz et al. (2000) showed that females of *H. versicolor* did not  
375 preferentially choose a conspecific call by itself over an identical alternative call that was paired with the  
376 call of a predator. Consistent with these findings, most females ( $\geq 70\%$ ) chose one of the four stimuli in  
377 our auditory streaming test (including the ABAB and AABB stimuli), which suggests B pulses were not  
378 inherently aversive. Results from an additional four-alternative choice test (see Supplementary  
379 Information) indicated B pulses can even be attractive in some stimulus contexts. Therefore, it seems  
380 highly unlikely that a perceived B–B– stream in the ABAB stimulus was in any way aversive in our test of  
381 auditory streaming. Finally, Stratman et al. (2021) demonstrated that females of *H. versicolor*  
382 preferentially approach small clusters of calling males over males calling in isolation. Had females  
383 perceptually segregated the ABAB stimuli into separate streams, one preferred (A–A–) and one less  
384 preferred (B–B–), then we might have expected the perceived presence of two males in close proximity to  
385 impart greater behavioral salience to the ABAB stimulus. Based on this other work, we interpret the lack  
386 of a significant preference for ABAB in our experiment as indicating that the pulse rise time differences  
387 did not promote auditory streaming.

388 Our study is the first investigation of the effects of pulse rise time differences on auditory  
389 streaming in a non-human animal. As such, our findings contribute to the existing knowledge on the effect  
390 of temporal differences on auditory streaming. While our study shows no effects of pulse rise time  
391 differences on auditory streaming, it would be worth testing the same hypothesis in other species, such  
392 as in some grasshoppers, which also use rise time as a behaviorally salient signal trait (Helversen, 1993).  
393 Besides our study, the only other investigations of the effect of amplitude rise time alone on segregation  
394 of sounds have been in humans. Similar to our study, Hartmann & Johnson (1991) tested the segregation  
395 of *sequential* sound elements and found rise time differences to be a weak facilitator of stream  
396 segregation. In that study, segregation of short (4 s) interleaved sequences of melodies (A and B) having  
397 different rise times was not any better than when melodies A and B had the same rise times. In contrast  
398 to our findings and those of Hartmann & Johnson (1991), Bregman et al. (1994a, b) demonstrated that  
399 rise time differences can facilitate segregation of sounds that occur simultaneously (as opposed to  
400 sequentially). In the studies by Bregman et al. (1994a, b), the discriminability of target tones in a multi-

401 tone complex was better when the target exhibited a sudden rise compared to the other tones in the  
402 complex. Bregman et al. (1994a, b) speculated that a sudden onset or change in amplitude of target  
403 tones may “reset” the pitch-analysis mechanisms, leading to the segregation of target tones from the  
404 complex.

405         The apparent inability of rise time differences to promote sequential stream segregation in our  
406 study and that by Hartmann & Johnson (1991) must be considered in light of a well-known phenomenon  
407 in auditory streaming known as the “build up” effect. During segregation of *sequential* sounds, the percept  
408 of two distinct streams does not arise instantaneously but instead builds up over several seconds after  
409 stimulus onset (Bregman, 1978; Anstis & Saida, 1985; Micheyl et al., 2005; Deike et al., 2012).  
410 Behavioral measurements in humans (Anstis & Saida, 1985; Bregman, 1978; Thompson et al., 2011),  
411 ferrets (Ma et al., 2010) and budgerigars (Cai et al., 2018) demonstrate that when hearing interleaved  
412 tone sequences that differ acoustically, subjects initially perceive a single stream. The probability of  
413 perceiving two streams increases as the sequence progresses. This build-up of a two-stream percept  
414 over time has been attributed to the long-term adaptation of neural responses, as demonstrated in  
415 mammals (Micheyl et al., 2005; Snyder et al., 2006; Pressnitzer et al., 2008) and songbirds (Bee et al.,  
416 2010). Importantly, previous studies on the build-up of auditory streaming used long interleaved  
417 sequences (> 10 s) and found that the build-up of a two-stream percept took several seconds (5-10 s). In  
418 contrast, the study by Hartmann & Johnson (1991), which failed to find strong evidence for sequential  
419 stream segregation based on differences in rise time, used an overall stimulus duration that was relatively  
420 short at 4 s. While our study involved similar ABAB interleaved sound sequences, our stimulus design  
421 was constrained by the requirement to stimulate natural communication signals. Consequently, one  
422 limitation of our study is that it only examined auditory streaming over relatively short sequences of pulses  
423 within calls that were < 1 s in duration. It is primarily for this reason that our main conclusion, namely that  
424 salient pulse rise time differences do not promote stream segregation in gray treefrogs, must remain  
425 provisional. The ability of perceptually salient differences in pulse rise time to impact auditory streaming  
426 using longer stimulus sequences remains to be investigated in frogs.

427         Previous investigations of perceptual organization in treefrogs illustrate the importance of  
428 considering both stimulus design and the perceptual task. For example, previous studies of *H.*  
429 *chrysoscelis* using short, call-like sequences of pulses similar to those used in the present study have  
430 revealed the importance of common onsets/offsets (Gupta & Bee, 2020) and common spatial location  
431 (Bee, 2010) in promoting simultaneous integration of the two harmonics in the pulses of gray treefrogs  
432 calls. In contrast to the study by Bee (2010), the effect of spatial separation between consecutive pulses  
433 in short, call-like pulse sequences had markedly less impact on promoting sequential segregation in both  
434 *H. versicolor* (Schwartz & Gerhardt, 1995 and Schwartz & Del Monte, 2019) and *H. chrysoscelis* (Bee &  
435 Riemersma, 2008). This discrepancy in the strength of spatial separation as a segregation cue across  
436 sequential versus simultaneous segregation tasks parallels the contrast between findings on pulse rise

437 time from the present study of frogs and those of humans by Bregman et al. (1994 a, b). One study of  
438 sequential segregation in *H. chrysoscelis* found that females could segregate a short, call-like sequence  
439 of pulses (A–A–) that was periodically interleaved with the pulses in a long (5 min) and continuous  
440 sequence of pulses (B–B–) differing in frequency, provided there was sufficient frequency separation  
441 between the A and B pulses (Nityananda & Bee, 2011). Whether pulse rise time differences might  
442 promote auditory streaming using a similar stimulus paradigm remains to be investigated.

443 Finally, it is also worth consider the lack of an effect of pulse rise time in the light of complex cue  
444 interactions during auditory streaming. In natural auditory scenes, multiple cues, or acoustic differences,  
445 are available to a receiver and may be differentially weighed during auditory streaming. For instance,  
446 Elhilali et al. (2009) tested auditory streaming in a cue conflict scenario using two sequences (A–A– and  
447 B–B–) that exhibited fairly large frequency separation, which promotes segregation, but shared coherent  
448 temporal onsets and offsets, which promotes integration. They found that coherent temporal  
449 onsets/offsets override frequency separation during auditory streaming as human subjects reported  
450 hearing a single stream (indicating integration). In other cases, different cues can also impact auditory  
451 streaming in an additive fashion. Micheyl et al. (2013), for example, found that inharmonicity (sounds  
452 having different fundamental frequencies) and temporal incoherence additively facilitate the segregation  
453 of sounds in humans. Importantly, while Hartmann & Johnson (1991) showed a weak effect of rise times  
454 on sound segregation (in the absence of frequency differences), Singh & Bregman (1997) showed an  
455 additive effect of rise times and frequency differences on stream segregation in humans. Non-human  
456 animals also incorporate cue interactions during auditory streaming, as seen for European starlings,  
457 Budgerigars and Zebra finches (Dent et al., 2016; Itatani & Klump, 2020). Additionally, there is also  
458 evidence for no interaction, as shown by Schwartz & Del Monte (2019) for spectral and spatial cues in  
459 *Hyla versicolor*. In the present study, A and B pulses had the same carrier frequencies, but different pulse  
460 rise times. It might be the case that spectral similarity overrides pulse rise time differences during auditory  
461 streaming. In such a case, we would expect that spectral similarity between A and B pulses promotes  
462 their integration irrespective of the differences in pulse rise times, which is in line with the findings of this  
463 study. Further, it might also be the case that differences in pulse rise time additively interacts with spectral  
464 differences to promote stream segregation. Additional studies that manipulate pulse rise time along with  
465 other potential cues, such as spectral differences or differences in spatial location, will be needed to  
466 uncover any interaction effects.

## 467 **Acknowledgements**

468 We thank Saumya Gupta, Katie Krueger, Sophie Barno, Kaitlyn Bonnema, Sarah Bonnema, Olivia Groth,  
469 Rishi Gulati, Collin Meyer and Annika Ruppert for their help collecting and testing frogs, and Mike  
470 Goodnature from Ramsey County Parks and Recreation for generous access to collection sites. We are  
471 also grateful to Marlene Zuk, Andrew Oxenham, Katie Krueger and Satyabhama de Oliveira for feedback  
472 on this manuscript.

473 **Author contributions** LK and MAB: designed the study. LK and SA: conducted experiments. LK and  
474 SA: analyzed data. LK, SA and MAB: prepared the manuscript.

475 **Funding** This research was funded in part by grants to MAB from the National Science Foundation (IOS-  
476 1452831 and IOS-2022253) and by grants and fellowships to LK from the University of Minnesota  
477 Graduate Program in Ecology, Evolution, and Behavior, the Bell Museum of Natural History, the  
478 University of Minnesota Graduate School and the Animal Behavior Society.

#### 479 **Declarations**

480 **Conflicts of interest** The authors declare no competing or financial interests.

481 **Ethical approval** This research was approved by the University of Minnesota Institutional Animal Care  
482 and Use Committee (#0602A-81890 and #2001-37746A).

#### 483 **References**

484 Akre, K. L., & Ryan, M. J. (2010). Proximity-dependent Response to Variably Complex Mating  
485 Signals in Túngara Frogs (*Physalaemus pustulosus*). *Ethology*, 116(12), 1138–1145.  
486 <https://doi.org/10.1111/j.1439-0310.2010.01825.x>

487 Anstis, S., & Saida, S. (1985). Adaptation to Auditory Streaming of Frequency-Modulated  
488 Tones. *Journal of Experimental Psychology: Human Perception and Performance*, 11(3),  
489 257–271. <https://doi.org/10.1037/0096-1523.11.3.257>

490 Bee, M. A. (2004). Equipment Review: Sound Ruler Acoustical Analysis: a free, open code,  
491 multi-platform sound analysis and graphing package. *Bioacoustics*, 14(2), 171–178.

492 Bee, M. A. (2010). Spectral preferences and the role of spatial coherence in simultaneous  
493 integration in gray treefrogs (*Hyla chrysoscelis*). *Journal of Comparative Psychology*,  
494 124(4), 412–424. <https://doi.org/10.1037/a0020307>

495 Bee, M. A. (2015). Treefrogs as animal models for research on auditory scene analysis and the  
496 cocktail party problem. *International Journal of Psychophysiology*, 95(2), 216–237.  
497 <https://doi.org/10.1016/j.ijpsycho.2014.01.004>

498 Bee, M. A., & Micheyl, C. (2008). The Cocktail Party Problem: What Is It? How Can It Be  
499 Solved? And Why Should Animal Behaviorists Study It? In *Journal of Comparative*  
500 *Psychology* (Vol. 122, Issue 3, pp. 235–251). American Psychological Association.  
501 <https://doi.org/10.1037/0735-7036.122.3.235>

502 Bee, M. A., Micheyl, C., Oxenham, A. J., & Klump, G. M. (2010). Neural adaptation to tone

- 503 sequences in the songbird forebrain: Patterns, determinants, and relation to the build-up of  
504 auditory streaming. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural,*  
505 *and Behavioral Physiology*, 196(8), 543–557. <https://doi.org/10.1007/s00359-010-0542-4>
- 506 Bee, M. A., & Riemersma, K. K. (2008). Does common spatial origin promote the auditory  
507 grouping of temporally separated signal elements in grey treefrogs? *Animal Behaviour*,  
508 76(3), 831–843. <https://doi.org/10.1016/j.anbehav.2008.01.026>
- 509 Blair, W. F. (1958). Call Difference as an Isolation Mechanism in Florida Species of Hylid Frogs.  
510 *Florida Acad Sci*, 21(1), 32–48.
- 511 Booker, W. W., Gerhardt, H. C., Lemmon, A. R., Ptacek, M. B., Hassinger, A. T. B., Schul, J., &  
512 Lemmon, E. M. (2022). The Complex History of Genome Duplication and Hybridization in  
513 North American Gray Treefrogs. *Molecular Biology and Evolution*, 39(2), msab316.  
514 <https://doi.org/10.1093/molbev/msab316>
- 515 Bregman, A. S. (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology:*  
516 *Human Perception and Performance*, 4(3), 380–387. [https://doi.org/10.1037/0096-](https://doi.org/10.1037/0096-1523.4.3.380)  
517 1523.4.3.380
- 518 Bregman, A. S., Ahad, P. A., & Kim, J. (1994). Resetting The Pitch-Analysis System. 2. Role Of  
519 Sudden Onsets And Offsets In The Perception Of Individual Components In A Cluster Of  
520 Overlapping Tones. *Journal of the Acoustical Society of America*, 96(5), 2694–2703.  
521 <https://doi.org/10.1121/1.411277>
- 522 Bregman, A. S., Ahad, P., Kim, J., & Melnerich, L. (1994). Resetting the pitch-analysis system:  
523 1. Effects of rise times of tones in noise backgrounds or of harmonics in a complex tone.  
524 *Perception & Psychophysics*, 56(2), 155–162. <https://doi.org/10.3758/BF03213894>
- 525 Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of  
526 order in rapid sequences of tones. *Journal of Experimental Psychology*, 89(2), 244–249.  
527 <https://doi.org/10.1037/h0031163>
- 528 Bush, S. L., Gerhardt, H. C., & Schul, J. (2002). Pattern recognition and call preferences in  
529 treefrogs (Anura: Hylidae): A quantitative analysis using a no-choice paradigm. *Animal*  
530 *Behaviour*, 63(1), 7–14. <https://doi.org/10.1006/anbe.2001.1880>
- 531 Cai, H., Screven, L. A., & Dent, M. L. (2018). Behavioral measurements of auditory streaming  
532 and build-up by budgerigars ( *Melopsittacus undulatus* ) . *The Journal of the Acoustical*

- 533            *Society of America*, 144(3), 1508–1516. <https://doi.org/10.1121/1.5054297>
- 534 Carl Gerhardt, H. (1978). Temperature coupling in the vocal communication system of the gray  
535 tree frog, *Hyla versicolor*. *Science*, 199(4332), 992–994.  
536 <https://doi.org/10.1126/science.199.4332.992>
- 537 Cherry, E. C., & Taylor, W. K. (1954). Some Further Experiments upon the Recognition of  
538 Speech, with One and with Two Ears. *Journal of the Acoustical Society of America*, 26(4),  
539 554–559. <https://doi.org/10.1121/1.1907373>
- 540 Christison-Lagay, K. L., & Cohen, Y. E. (2014). Behavioral correlates of auditory streaming in  
541 rhesus macaques. *Hearing Research*, 309, 17–25.  
542 <https://doi.org/10.1016/j.heares.2013.11.001>
- 543 Darwin, Chris J. (1997). Auditory grouping. *Trends in Cognitive Sciences*, 1(9), 327–333.  
544 [https://doi.org/10.1016/S1364-6613\(97\)01097-8](https://doi.org/10.1016/S1364-6613(97)01097-8)
- 545 Darwin, Christopher J. (2007). Spatial Hearing and Perceiving Sources. In *Auditory Perception*  
546 *of Sound Sources* (pp. 215–232). Springer. [https://doi.org/10.1007/978-0-387-71305-2\\_8](https://doi.org/10.1007/978-0-387-71305-2_8)
- 547 Deike, S., Heil, P., Böckmann-Barthel, M., & Brechmann, A. (2012). The build-up of auditory  
548 stream segregation: A different perspective. In *Frontiers in Psychology* (Vol. 3, Issue OCT).  
549 <https://doi.org/10.3389/fpsyg.2012.00461>
- 550 Dent, M. L., & Bee, M. A. (2018). *Principles of Auditory Object Formation by Nonhuman Animals*  
551 (pp. 47–82). Springer, New York, NY. [https://doi.org/10.1007/978-1-4939-8574-6\\_3](https://doi.org/10.1007/978-1-4939-8574-6_3)
- 552 Dent, M. L., Martin, A. K., Flaherty, M. M., & Neilans, E. G. (2016). Cues for auditory stream  
553 segregation of birdsong in budgerigars and zebra finches: Effects of location, timing,  
554 amplitude, and frequency. *The Journal of the Acoustical Society of America*, 139(2), 674–  
555 683. <https://doi.org/10.1121/1.4941322>
- 556 Diekamp, B., & Gerhardt, H. C. (1995). Selective phonotaxis to advertisement calls in the gray  
557 treefrog *Hyla versicolor*: behavioral experiments and neurophysiological correlates. *Journal*  
558 *of Comparative Physiology A*, 177(2), 173–190. <https://doi.org/10.1007/BF00225097>
- 559 Dowling, W. J. (2012). Music perception. *Oxford Handbook of Auditory Science Hearing*, 12(7),  
560 4473–4482. <https://doi.org/10.1093/oxfordhb/9780199233557.013.0010>
- 561 Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J., & Shamma, S. A. (2009). Temporal Coherence

- 562 in the Perceptual Organization and Cortical Representation of Auditory Scenes. *Neuron*,  
563 61(2), 317–329. <https://doi.org/10.1016/j.neuron.2008.12.005>
- 564 Farris, H. E., Rand, A. S., & Ryan, M. J. (2005). The effects of time, space and spectrum on  
565 auditory grouping in túngara frogs. *Journal of Comparative Physiology. A, Neuroethology,*  
566 *Sensory, Neural, and Behavioral Physiology*, 191(12), 1173–1183.  
567 <https://doi.org/10.1007/s00359-005-0041-1>
- 568 Farris, Hamilton E., Rand, A. S., & Ryan, M. J. (2002). The effects of spatially separated call  
569 components on phonotaxis in túngara frogs: Evidence for auditory grouping. *Brain,*  
570 *Behavior and Evolution*, 60(3), 181–188. <https://doi.org/10.1159/000065937>
- 571 Fay, R. R. (2000). Spectral contrasts underlying auditory stream segregation in goldfish  
572 (*Carassius auratus*). *JARO - Journal of the Association for Research in Otolaryngology*,  
573 1(2), 120–128. <https://doi.org/10.1007/s101620010015>
- 574 Fay, Richard R. (1998). Auditory stream segregation in goldfish (*Carassius auratus*). *Hearing*  
575 *Research*, 120(1–2), 69–76. [https://doi.org/10.1016/S0378-5955\(98\)00058-6](https://doi.org/10.1016/S0378-5955(98)00058-6)
- 576 Fishbein, A. R., Idsardi, W. J., Ball, G. F., & Dooling, R. J. (2020). Sound sequences in  
577 birdsong: How much do birds really care? *Philosophical Transactions of the Royal Society*  
578 *B: Biological Sciences*, 375(1789), 20190044. <https://doi.org/10.1098/rstb.2019.0044>
- 579 Gerhardt, H. C., & Schul, J. (1999). A quantitative analysis of behavioral selectivity for pulse  
580 rise-time in the gray treefrog, *Hyla versicolor*. *Journal of Comparative Physiology - A*  
581 *Sensory, Neural, and Behavioral Physiology*, 185(1), 33–40.  
582 <https://doi.org/10.1007/s003590050363>
- 583 Gerhardt, H. Carl. (1975). Sound pressure levels and radiation patterns of the vocalizations of  
584 some North American frogs and toads. *Journal of Comparative Physiology ■ A*, 102(1), 1–  
585 12. <https://doi.org/10.1007/BF00657481>
- 586 Gerhardt, H. Carl. (1981). Mating call recognition in the green treefrog (*Hyla cinerea*):  
587 Importance of two frequency bands as a function of sound pressure level. *Journal of*  
588 *Comparative Physiology □ A*, 144(1), 9–16. <https://doi.org/10.1007/BF00612792>
- 589 Gerhardt, H. Carl. (1987). Evolutionary and neurobiological implications of selective phonotaxis  
590 in the green treefrog, *Hyla cinerea*. *Animal Behaviour*, 35(5), 1479–1489.  
591 [https://doi.org/10.1016/S0003-3472\(87\)80020-9](https://doi.org/10.1016/S0003-3472(87)80020-9)



- 592 Gerhardt, H. Carl. (2001). Acoustic communication in two groups of closely related treefrogs. In  
593 *Advances in the Study of Behavior* (Vol. 30, pp. 99–167). Academic Press.  
594 [https://doi.org/10.1016/s0065-3454\(01\)80006-1](https://doi.org/10.1016/s0065-3454(01)80006-1)
- 595 Gerhardt, H. Carl. (2005). Acoustic spectral preferences in two cryptic species of grey treefrogs:  
596 Implications for mate choice and sensory mechanisms. *Animal Behaviour*, 70(1), 39–48.  
597 <https://doi.org/10.1016/j.anbehav.2004.09.021>
- 598 Gerhardt, H. Carl. (2005b). Advertisement-call preferences in diploid-tetraploid treefrogs (*Hyla*  
599 *chrysoscelis* and *Hyla versicolor*): Implications for mate choice and the evolution of  
600 communication systems. *Evolution*, 59(2), 395–408. [https://doi.org/10.1111/j.0014-](https://doi.org/10.1111/j.0014-3820.2005.tb00998.x)  
601 [3820.2005.tb00998.x](https://doi.org/10.1111/j.0014-3820.2005.tb00998.x)
- 602 Gerhardt, H. Carl. (2008). Phonotactic selectivity in two cryptic species of gray treefrogs: Effects  
603 of differences in pulse rate, carrier frequency and playback level. *Journal of Experimental*  
604 *Biology*, 211(16), 2609–2616. <https://doi.org/10.1242/jeb.019612>
- 605 Gerhardt, H. Carl. (1995). Phonotaxis in Female Frogs and Toads: Execution and Design of  
606 Experiments. *Methods in Comparative Psychoacoustics*, 209–220.  
607 [https://doi.org/10.1007/978-3-0348-7463-2\\_18](https://doi.org/10.1007/978-3-0348-7463-2_18)
- 608 Gerhardt, H. Carl, & Doherty, J. A. (1988). Acoustic communication in the gray treefrog, *Hyla*  
609 *versicolor*: evolutionary and neurobiological implications. *Journal of Comparative*  
610 *Physiology A*, 162(2), 261–278. <https://doi.org/10.1007/BF00606090>
- 611 Gerhardt, H. Carl, Dyson, M. L., Tanner, S. D., & Murphy, C. G. (1994). Female treefrogs do not  
612 avoid heterospecific calls as they approach conspecific calls: Implications for mechanisms  
613 of mate choice. *Animal Behaviour*, 47(6), 1323–1332.  
614 <https://doi.org/10.1006/anbe.1994.1180>
- 615 Gerhardt, H. Carl, Martínez-Rivera, C. C., Schwartz, J. J., Marshall, V. T., & Murphy, C. G.  
616 (2007). Preferences based on spectral differences in acoustic signals in four species of  
617 treefrogs (Anura: Hylidae). *Journal of Experimental Biology*, 210(17), 2990–2998.  
618 <https://doi.org/10.1242/jeb.006312>
- 619 Greenfield, M. D. (2005). Mechanisms and Evolution of Communal Sexual Displays in  
620 Arthropods and Anurans. In *Advances in the Study of Behavior* (Vol. 35, pp. 1–62).  
621 [https://doi.org/10.1016/S0065-3454\(05\)35001-7](https://doi.org/10.1016/S0065-3454(05)35001-7)

- 622 Gridi-Papp, M. (2007). Sound ruler. *Acoustic Analysis. Version 0.9*, 6.
- 623 Gupta, S., Alluri, R. K., Rose, G. J., & Bee, M. A. (2021). Neural basis of acoustic species  
624 recognition in a cryptic species complex. *Journal of Experimental Biology*, 224(23),  
625 jeb243405. <https://doi.org/10.1242/jeb.243405>
- 626 Gupta, S., & Bee, M. A. (2020). Treefrogs exploit temporal coherence to form perceptual objects  
627 of communication signals. *Biology Letters*, 16(9), 20200573.  
628 <https://doi.org/10.1098/rsbl.2020.0573>
- 629 Gupta, S., & Bee, M. A. (2023). Female preferences for the spectral content of advertisement  
630 calls in Cope's gray treefrog (*Hyla chrysoscelis*). *Journal of Comparative Physiology A:  
631 Neuroethology, Sensory, Neural, and Behavioral Physiology*, 209(1), 31–45.  
632 <https://doi.org/10.1007/s00359-022-01575-w>
- 633 Hartmann, W. M., & Johnson, D. (1991). Stream Segregation and Peripheral Channeling. *Music  
634 Perception*, 9(2), 155–183. <https://doi.org/10.2307/40285527>
- 635 Hulse, S. H. (2002). Auditory scene analysis in animal communication. *Advances in the Study of  
636 Behavior*, 31, 163–200. [https://doi.org/10.1016/s0065-3454\(02\)80008-0](https://doi.org/10.1016/s0065-3454(02)80008-0)
- 637 Hulse, S. H., MacDougall-Shackleton, S. A., & Wisniewski, A. B. (1997). Auditory Scene  
638 Analysis by Songbirds: Stream Segregation of Birdsong by European Starlings (*Sturnus  
639 vulgaris*). *Journal of Comparative Psychology*, 111(1), 3–13. [https://doi.org/10.1037/0735-  
640 7036.111.1.3](https://doi.org/10.1037/0735-7036.111.1.3)
- 641 Hyland Bruno, J., & Tchernichovski, O. (2019). Regularities in zebra finch song beyond the  
642 repeated motif. *Behavioural Processes*, 163, 53–59.  
643 <https://doi.org/10.1016/j.beproc.2017.11.001>
- 644 Itatani, N., & Klump, G. M. (2009). Auditory streaming of amplitude-modulated sounds in the  
645 songbird forebrain. *Journal of Neurophysiology*, 101(6), 3212–3225.  
646 <https://doi.org/10.1152/jn.91333.2008>
- 647 Itatani, N., & Klump, G. M. (2014). Neural correlates of auditory streaming in an objective  
648 behavioral task. *Proceedings of the National Academy of Sciences of the United States of  
649 America*, 111(29), 10738–10743. <https://doi.org/10.1073/pnas.1321487111>
- 650 Itatani, N., & Klump, G. M. (2020). Interaction of spatial and non-spatial cues in auditory stream  
651 segregation in the European starling. *European Journal of Neuroscience*, 51(5), 1191–

- 652 1200. <https://doi.org/10.1111/ejn.13716>
- 653 Izumi, A. (2002). Auditory stream segregation in Japanese monkeys. *Cognition*, 82(3), B113–  
654 B122. [https://doi.org/10.1016/S0010-0277\(01\)00161-5](https://doi.org/10.1016/S0010-0277(01)00161-5)
- 655 Ma, L., Micheyl, C., Yin, P., Oxenham, A. J., & Shamma, S. A. (2010). Behavioral measures of  
656 auditory streaming in ferrets (*Mustela putorius*). *Journal of Comparative Psychology*,  
657 124(3), 317–330. <https://doi.org/10.1037/a0018273>
- 658 MacDougall-Shackleton, S. A., Hulse, S. H., & Ball, G. F. (1998). Neural bases of song  
659 preferences in female zebra finches (*Taeniopygia guttata*). *NeuroReport*, 9(13), 3047–  
660 3052. <https://doi.org/10.1097/00001756-199809140-00024>
- 661 McDermott, J. H. (2009). The cocktail party problem. *Current Biology*, 19(22), R1024–R1027.  
662 <https://doi.org/10.1016/j.cub.2009.09.005>
- 663 McDermott, J. H., & Oxenham, A. J. (2008). Music perception, pitch, and the auditory system.  
664 *Current Opinion in Neurobiology*, 18(4), 452–463.
- 665 Mercado, E., Herman, L. M., & Pack, A. A. (2005). Stereotypical sound patterns in humpback  
666 whale songs: Usage and function. *Aquatic Mammals*, 29(1), 37–52.  
667 <https://doi.org/10.1578/016754203101024068>
- 668 Micheyl, C., Kreft, H., Shamma, S., & Oxenham, A. J. (2013). Temporal coherence versus  
669 harmonicity in auditory stream formation. *The Journal of the Acoustical Society of America*,  
670 133(3), EL188–EL194. <https://doi.org/10.1121/1.4789866>
- 671 Micheyl, C., & Oxenham, A. J. (2010). Pitch, harmonicity and concurrent sound segregation:  
672 Psychoacoustical and neurophysiological findings. *Hearing Research*, 266(1–2), 36–51.  
673 <https://doi.org/10.1016/j.heares.2009.09.012>
- 674 Micheyl, C., Tian, B., Carlyon, R. P., & Rauschecker, J. P. (2005). Perceptual organization of  
675 tone sequences in the auditory cortex of awake macaques. *Neuron*, 48(1), 139–148.  
676 <https://doi.org/10.1016/j.neuron.2005.08.039>
- 677 Middlebrooks, J. C., & Bremen, P. (2013). Spatial stream segregation by auditory cortical  
678 neurons. *Journal of Neuroscience*, 33(27), 10986–11001.  
679 <https://doi.org/10.1523/JNEUROSCI.1065-13.2013>
- 680 Moore, B. C. J., & Gockel, H. E. (2012). Properties of auditory stream formation. In

- 681 *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 367, Issue  
682 1591, pp. 919–931). <https://doi.org/10.1098/rstb.2011.0355>
- 683 Narins, P. M. (2003). Acoustic Communication in Insects and Anurans. Common Problems and  
684 Diverse Solutions. In *Copeia* (Vol. 2003, Issue 3). University of Chicago Press.  
685 <https://doi.org/10.1643/ot-03-039>
- 686 Nelken, I. (1997). Demonstrations of auditory scene analysis: The perceptual organization of  
687 sound. In *Trends in Neurosciences* (Vol. 20, Issue 4). MIT Press.  
688 [https://doi.org/10.1016/S0166-2236\(96\)01022-3](https://doi.org/10.1016/S0166-2236(96)01022-3)
- 689 Nihoul, J. C. J. (2004). Echolocation in Bats and Dolphins. In *Journal of Marine Systems* (Vol.  
690 50, Issues 3–4). University of Chicago press. <https://doi.org/10.1016/j.jmarsys.2004.01.009>
- 691 Nityananda, V., & Bee, M. A. (2011). Finding your mate at a cocktail party: Frequency  
692 separation promotes auditory stream segregation of concurrent voices in multi-species frog  
693 choruses. *PLoS ONE*, 6(6). <https://doi.org/10.1371/journal.pone.0021191>
- 694 Noble, G. K., & Hassler, W. G. (1936). Three Salientia of Geographic Interest from Southern  
695 Maryland. *Copeia*, 1936(1), 63. <https://doi.org/10.2307/1436381>
- 696 Noda, T., Kanzaki, R., & Takahashi, H. (2013). Stimulus phase locking of cortical oscillation for  
697 auditory stream segregation in rats. *PLoS ONE*, 8(12), e83544.  
698 <https://doi.org/10.1371/journal.pone.0083544>
- 699 Platz, J. E., & Forester, D. C. (1988). Geographic Variation in Mating Call among the Four  
700 Subspecies of the Chorus Frog: *Pseudacris triseriata* (Wied). *Copeia*, 1988(4), 1062.  
701 <https://doi.org/10.2307/1445734>
- 702 Pressnitzer, D., Sayles, M., Micheyl, C., & Winter, I. M. (2008). Perceptual Organization of  
703 Sound Begins in the Auditory Periphery. *Current Biology*, 18(15), 1124–1128.  
704 <https://doi.org/10.1016/j.cub.2008.06.053>
- 705 Prestwich, K. N. (1994). The energetics of acoustic signaling in anurans and insects. *Integrative  
706 and Comparative Biology*, 34(6), 625–643. <https://doi.org/10.1093/icb/34.6.625>
- 707 Ptacek, M. B., Gerhardt, H. C., & Sage, R. D. (1994). Speciation by Polyploidy in Treefrogs:  
708 Multiple Origins of the Tetraploid, *Hyla versicolor*. *Evolution*, 48(3), 898.  
709 <https://doi.org/10.2307/2410495>

- 710 Remez, R. E. (2021). Perceptual organization of speech. *The Handbook of Speech Perception*,  
711 3–27. <https://doi.org/10.1002/9781119184096.ch1>
- 712 Repp, B. H. (1988). Integration and segregation in speech perception. *Language and Speech*,  
713 31(3), 239–271. <https://doi.org/10.1177/002383098803100302>
- 714 Schul, J., & Sheridan, R. A. (2006). Auditory stream segregation in an insect. *Neuroscience*,  
715 138(1), 1–4. <https://doi.org/10.1016/j.neuroscience.2005.11.023>
- 716 Schul, Johannes, & Bush, S. L. (2002). Non-parallel coevolution of sender and receiver in the  
717 acoustic communication system of treefrogs. *Proceedings of the Royal Society B:*  
718 *Biological Sciences*, 269(1502), 1847–1852. <https://doi.org/10.1098/rspb.2002.2092>
- 719 Schwartz, J J, & Gerhardt, H. C. (1995). Directionality of the auditory system and call pattern  
720 recognition during acoustic interference in the gray treefrog, *Hyla versicolor*. *Auditory*  
721 *Neuroscience*, 1, 195–206.
- 722 Schwartz, Joshua J., Bee, M. A., & Tanner, S. D. (2000). A behavioral and neurobiological study  
723 of the responses of gray treefrogs, *Hyla versicolor*, to the calls of a predator, *Rana*  
724 *catesbeiana*. *Herpetologica*, 56(1), 27–37.
- 725 Schwartz, Joshua J., Buchanan, B. W., & Gerhardt, H. C. (2002). Acoustic interactions among  
726 male gray treefrogs, *Hyla versicolor*, in a chorus setting. *Behavioral Ecology and*  
727 *Sociobiology*, 53(1), 9–19. <https://doi.org/10.1007/s00265-002-0542-7>
- 728 Schwartz, Joshua J., & Marshall, V. T. (2006). Forms of call overlap and their impact on  
729 advertisement call attractiveness to females of the gray treefrog, *Hyla versicolor*.  
730 *Bioacoustics*, 16(1), 39–56. <https://doi.org/10.1080/09524622.2006.9753563>
- 731 Schwartz, Joshua J., & Serratto Del Monte, M. E. (2019). Spatially-mediated call pattern  
732 recognition and the cocktail party problem in treefrog choruses: can call frequency  
733 differences help during signal overlap? *Bioacoustics*, 28(4), 312–328.  
734 <https://doi.org/10.1080/09524622.2018.1443836>
- 735 Shamma, S. A., Elhilali, M., & Micheyl, C. (2011). Temporal coherence and attention in auditory  
736 scene analysis. In *Trends in Neurosciences* (Vol. 34, Issue 3, pp. 114–123).  
737 <https://doi.org/10.1016/j.tins.2010.11.002>
- 738 Singh, P. G., & Bregman, A. S. (1997). The influence of different timbre attributes on the  
739 perceptual segregation of complex-tone sequences. *The Journal of the Acoustical Society*

- 740 *of America*, 102(4), 1943–1952. <https://doi.org/10.1121/1.419688>
- 741 Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric correlates of  
742 auditory stream segregation. *Journal of Cognitive Neuroscience*, 18(1), 1–13.  
743 <https://doi.org/10.1162/089892906775250021>
- 744 Stratman, K. D., Oldehoeft, E. A., & Höbel, G. (2021). Woe is the loner: Female treefrogs prefer  
745 clusters of displaying males over single “hotshot” males. *Evolution*, 75(12), 3026–3036.  
746 <https://doi.org/10.1111/evo.14376>
- 747 Thompson, S. K., Carlyon, R. P., & Cusack, R. (2011). An Objective Measurement of the Build-  
748 Up of Auditory Streaming and of Its Modulation by Attention. *Journal of Experimental*  
749 *Psychology: Human Perception and Performance*, 37(4), 1253–1262.  
750 <https://doi.org/10.1037/a0021925>
- 751 van-Noorden, L. P. A. S. (1974). Temporal Coherence in Random Tone Sequences. *IPO*  
752 *Annual Progress Report*, 9, 4–21.
- 753 von Helversen, D. (1993). “Absolute steepness” of ramps as an essential cue for auditory  
754 pattern recognition by a grasshopper (Orthoptera; Acrididae; Chorthippus biguttulus L.).  
755 *Journal of Comparative Physiology A*, 172(5), 633–639.  
756 <https://doi.org/10.1007/BF00213685>
- 757 von Helversen, Dagmar. (1984). Parallel processing in auditory pattern recognition and  
758 directional analysis by the grasshopper Chorthippus biguttulus L. (Acrididae). *Journal of*  
759 *Comparative Physiology A*, 154(6), 837–846. <https://doi.org/10.1007/BF00610684>
- 760 Ward, J. L., Love, E. K., Vélez, A., Buerkle, N. P., O’Bryan, L. R., & Bee, M. A. (2013).  
761 Multitasking males and multiplicative females: Dynamic signalling and receiver preferences  
762 in Cope’s grey treefrog. *Animal Behaviour*, 86(2), 231–243.  
763 <https://doi.org/10.1016/j.anbehav.2013.05.016>
- 764 Weber, T., & Thorson, J. (1988). Auditory behavior of the cricket - IV. Interaction of direction of  
765 tracking with perceived temporal pattern in split-song paradigms. *Journal of Comparative*  
766 *Physiology A*, 163(1), 13–22. <https://doi.org/10.1007/BF00611992>
- 767 Winn, H. E., Thompson, T. J., Cummings, W. C., Hain, J., Hudnall, J., Hays, H., & Steiner, W.  
768 W. (1981). Song of the humpback whale - Population comparisons. *Behavioral Ecology*  
769 *and Sociobiology*, 8(1), 41–46. <https://doi.org/10.1007/BF00302842>

770 Yao, J. D., Bremen, P., & Middlebrooks, J. C. (2015). Emergence of spatial stream segregation  
771 in the ascending auditory pathway. *Journal of Neuroscience*, 35(49), 16199–16212.  
772 <https://doi.org/10.1523/JNEUROSCI.3116-15.2015>

773

774