Supplementary material

Methods

(a) Recording and acoustic analyses

We used models of different predators, both occurring at the study site, to prompt aerial and mobbing alarm calls. Aerial calls were prompted with a gliding, life-size model adult collared sparrowhawk, *Accipiter cirrhocephalus* (details in [1, 2]). The model was made of foam, with components cut to shape with a hot wire, so that it could glide for about 15-20 m. One person threw the model from a distance of about 10 m from the focal bird, aimed to pass by the bird at a distance of about 5 m. The second person recorded alarm calls with a Marantz PMD661 audio recorder sampling wave files at 48 kHz and 24 bits, using a Sennheiser ME66 directional microphone. The mean distance from the calling bird was 8.2 m (range 5.4–11.5 m; recorded with a laser rangefinder). One aerial call was prompted when an Australian magpie, *Gymnorhina tibicen*, flew by, and recorded from a distance of 8.0 m. We used taxidermic models of the boobook owls, *Ninox novaeseelandiae*, to prompt mobbing alarm calls. One was stationary and the other was enhanced with animation. The animated model’s head was connected to a servo motor controlled by an Arduino circuit board (Arduino Micro; http://www.arduino.cc), so that it rotated up to 180 degrees in a series of pre-programmed movements. This was designed to make the model more realistic [3], although we found that both models prompted responses. Playbacks were used to attract fairy-wrens and so prompt approach, but turned off during recording. Playbacks included white-browed scrubwren, *Sericornis frontalis*, mobbing calls and fairy-wren territorial or mobbing calls, and delivered from a speaker placed next to the model. Calls were recorded with the same recorder and settings, but using the a Sennheiser ME62 omnidirectional microphone placed on a tripod 25 cm above the ground and 0.25–0.5 m from the model. Observers remained 10-15 m away from the model, and recorded calls for up to 5 min.

We selected high-quality recordings and measured their acoustic attributes in Raven Pro 1.5 [4] (**Figure S1**). All recordings of elements were scored for quality based on the amount of reverberation and overlapping sound, and the three best from each territory selected for analysis. A random sample was taken if there were more elements of similar quality from a territory, and poor-quality recordings were excluded. This resulted in 48 mobbing elements from 16 territories (3 per territory) and 32 aerial elements from 16 territories (9 territories had fewer than three high-quality elements). These recordings were gathered from 25 territories in total, with both aerial and mobbing calls recorded from seven. Recordings were then normalized to a constant peak amplitude and visualized in spectrograms using a Blackman window function, and a grid resolution of 0.208 ms and 188 Hz, with 95% overlap. We then placed a selection box around each element, using a fixed display scale for time and frequency, and measured 12 acoustic features designed to identify variation between call types. These features are listed in the main text.

1. Call synthesis

The accuracy of call synthesis was assessed using both acoustic measurements and spectral cross-correlation. The 12 acoustic features used to compare aerial and mobbing elements (Methods, main text) were also measured for the 32 synthetic aerial and mobbing elements (16 of each) used in playback experiments. Given that call synthesis was only possible in 16-bit, we re-measured the natural calls after converting the sound files to 16-bit. We then used a Principle Component Analysis in R version 3.4.3 [5] to compare the synthetic calls with the 16 aerial and mobbing elements on which they were based. We also compared all individual synthetic and natural elements using spectral cross-correlation in Raven Pro 1.5 [4]. This analysis estimated how closely each synthetic element matched the element on which it was based and all other natural elements of that alarm call type.

Synthetic aerial and mobbing calls were a close representation of natural calls, even to the extent of capturing variation among individual elements. Principle component analysis showed that synthetic calls clustered with their natural counterparts (**Figure S2, Table S1**). Furthermore, spectrographic cross-correlation showed that all but one of the 32 synthetic elements was more similar to the natural element on which it was based than any other element (**Table S2**). Mean cross-correlations between synthetic and matching natural elements were high for both aerial and mobbing elements (mean ± SD, range: aerial 0.89 ± 0.05, 0.78–0.95; mobbing 0.77 ± 0.13, 0.52–0.92). Cross-correlations were much lower for non-matching elements (aerial 0.44 ± 0.16, 0.03–0.78; mobbing 0.31 ± 0.12, 0.03–0.64).

(c) Playback experiments

All playback experiments used calls consisting of two elements at natural intervals and amplitude. Calls were made by repeating a single element after an interval of 250 ms, which was between the two means and within the natural range of both types (mean interval ± SD, range, n intervals: mobbing 463 ± 349, 143-1870 ms, n = 42; aerial 120 ± 125, 25–392 ms, n = 23). Both aerial and mobbing calls are commonly composed of two elements [6], and all multi-element, natural aerial calls provoke flight to cover [7]. Calls were broadcast from a Peerless 810921 tweeter speaker (frequency response 2–11 kHz) connected to a custom amplifier and Roland R-05 digital player, all mounted on the observer’s waist to allow a mobile playback system [2]. Playbacks were calibrated so that all call elements were broadcast at 50 dB at 10 m, a natural amplitude for each call type at that distance (mean ± SD, range, n elements: aerial = 52.7 ± 6.4, 37.1-64.5 dB, 42; mobbing = 45.7 ± 5.9, 34.2-59.4 dB, 138) and the average distance to focal birds.

All four playback experiments had a matched design, with a focal individual on all 16 territories receiving a unique set of all treatments within an experiment. Treatment order was balanced, with different alarm variants occurring equally frequently at each order in all experiments, and with an equal number of each sex selected as focal individuals. In Experiments 3 and 4, playback on a territory was always to the same focal individual. In Experiments 1 and 2, time constraints meant that an individual of the same sex and plumage status was occasionally substituted for the focal individual (5/16 and 4/16 territories respectively). Only males that were in blue plumage were included in the study, because plumage colour can affect response to alarm calls and most males are blue during the breeding season [8]. Focal individuals were observed for a minimum of 5 min, without any alarm calls or predators nearby, before each playback. Playbacks were carried out when the focal individual was 8–10 m from the speaker and foraging on the ground and 0.5–5.0 m from cover.

References

[1] Magrath, R. D., Haff, T. M., McLachlan, J. R. & Igic, B. 2015 Wild birds learn to eavesdrop on heterospecific alarm calls. *Curr. Biol.* **25**, 1-4.

[2] Magrath, R. D., Pitcher, B. J. & Gardner, J. L. 2007 A mutual understanding? Interspecific responses by birds to each other's aerial alarm calls. *Behav. Ecol.* **18**, 944-951.

[3] Billings, A. C., Greene, E. & MacArthur-Waltz, D. 2017 Steller’s jays assess and communicate about predator risk using detection cues and identity. *Behav. Ecol.* **28**, 776-783.

[4] Charif, R. A., Waack, A. M. & Strickman, L. M. 2010 *Raven Pro 1.4 User's Manual*. Ithaca, NY, Cornell Laboratory of Ornithology.

[5] R Core Team. 2018 R: A Language and Environment for Statistical Computing. Vienna, R Foundation for Statstical Computing.

[6] Rowley, I. & Russell, E. 1997 *Fairy-wrens and Grasswrens: Maluridae*. Oxford, Oxford University Press.

[7] Fallow, P. M. & Magrath, R. D. 2010 Eavesdropping on other species: mutual interspecific understanding of urgency information in avian alarm calls. *Anim. Behav.* **79**, 411-417.

[8] McQueen, A., Naimo, A. C., Teunissen, N., Magrath, R. D., Delhey, K. & Peters, A. 2017 Bright birds are cautious: seasonally conspicuous plumage prompts risk avoidance by male superb fairy-wrens. *Proc. R. Soc. B* **284**, 20170446.

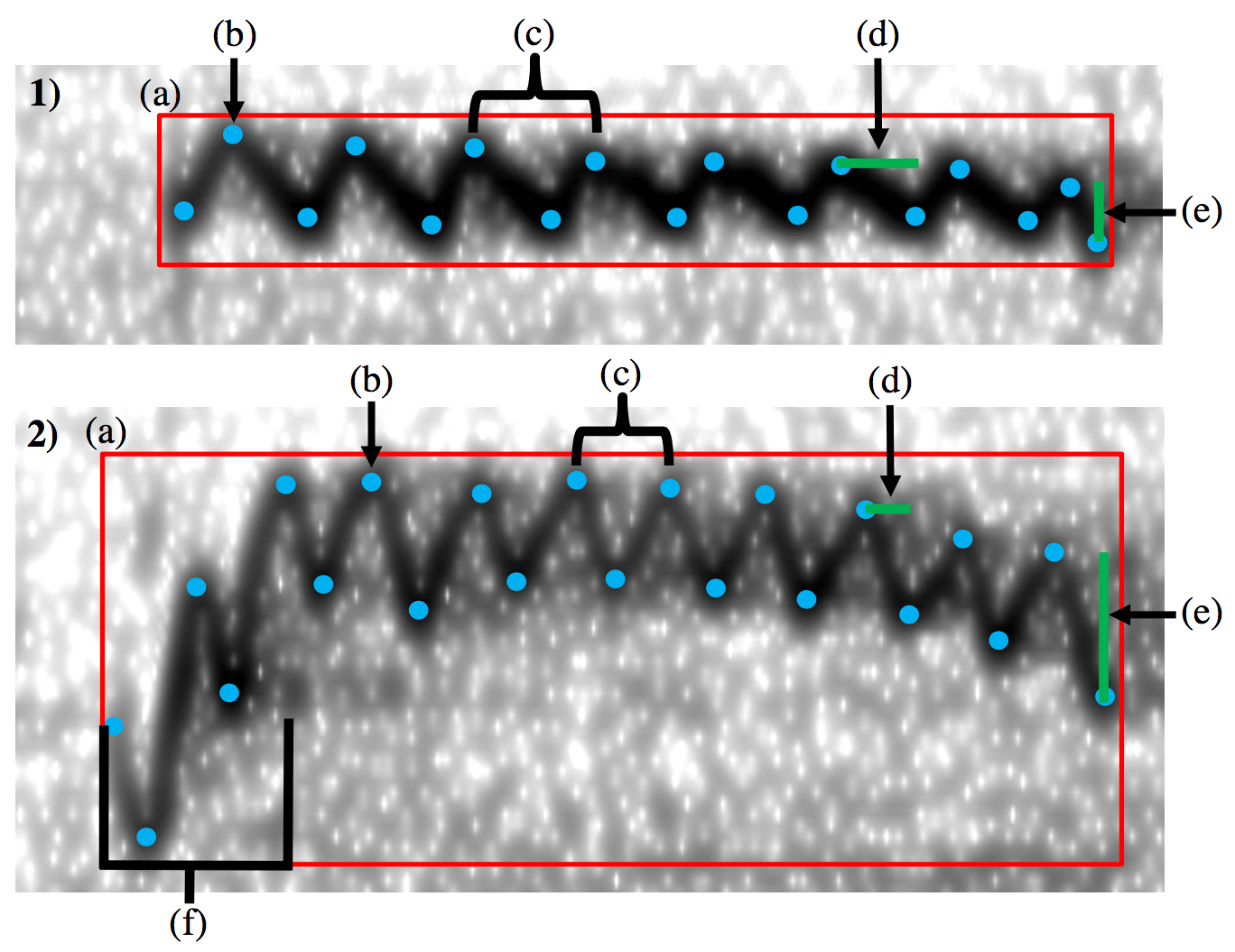
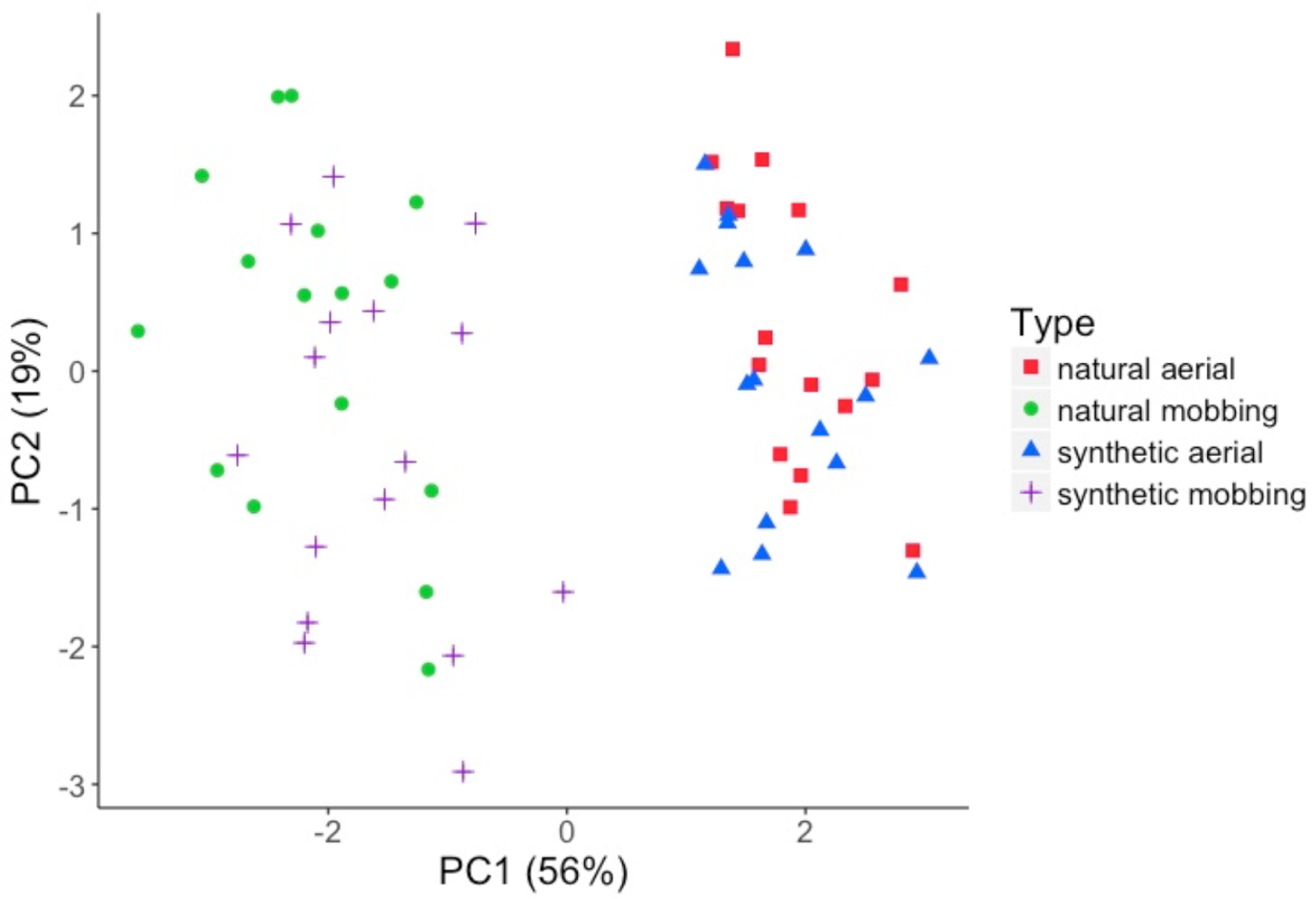


Figure S1. Measurements of acoustic features in Raven 1.5 Pro. Spectrograms of fairy-wren (1) aerial and (2) mobbing alarm call elements showing: (a) manual selection box; (b) inflection points; and (c) one full frequency cycle. The difference in (d) time and (e) frequency between successive inflection points were used to calculate time and frequency irregularity. The (f) hook component of the mobbing call is also indicated; measurement of the ‘non-hook’ component of the element entailed placing a separate manual selection box around the element but excluding this hook component.



**Figure S2**. Principal components plot of the acoustic features of natural and synthetic versions of aerial and mobbing alarm calls used in playback experiments (n = 16 for each type). **Table S1** presents loadings for PC1 and PC2.

Table S1. PC1 and PC2 loadings from a principal component analysis examining acoustic features of natural and synthetic versions of aerial and mobbing calls used in playback experiments (n = 16 examples of each type). Figure S2 shows the scatterplot.

|  |  |  |
| --- | --- | --- |
| Acoustic Feature | PC1 | PC2 |
| Duration (ms) | -0.110 | 0.479 |
| Low Frequency (Hz) | 0.491 | -0.048 |
| High Frequency (Hz) | 0.240 | -0.599 |
| Peak Frequency (Hz) | 0.475 | -0.078 |
| Slope (Hz) | -0.443 | -0.158 |
| Frequency Irregularity (Hz) | -0.472 | -0.125 |
| Rate Frequency Modulation (Hz) | -0.211 | -0.602 |

Table S2. Spectrographic cross-correlation for synthetic and natural versions of (a) aerial and (b) mobbing call elements. The column and row labels identify individual birds within each call type. The values are peak correlations, which have a maximum possible value of 1 (identical). Spectrographic correlations were carried out in Raven 1.5 Pro, using normalized sound files. The data are colour-coded, with elements that are more correlated in red and less correlated in blue. The synthetic version of the mobbing element of Bird 6 was the only one that was not most similar to its natural counterpart out of all 32 synthetic calls. The Supplemental Methods text gives means ± SD and ranges.

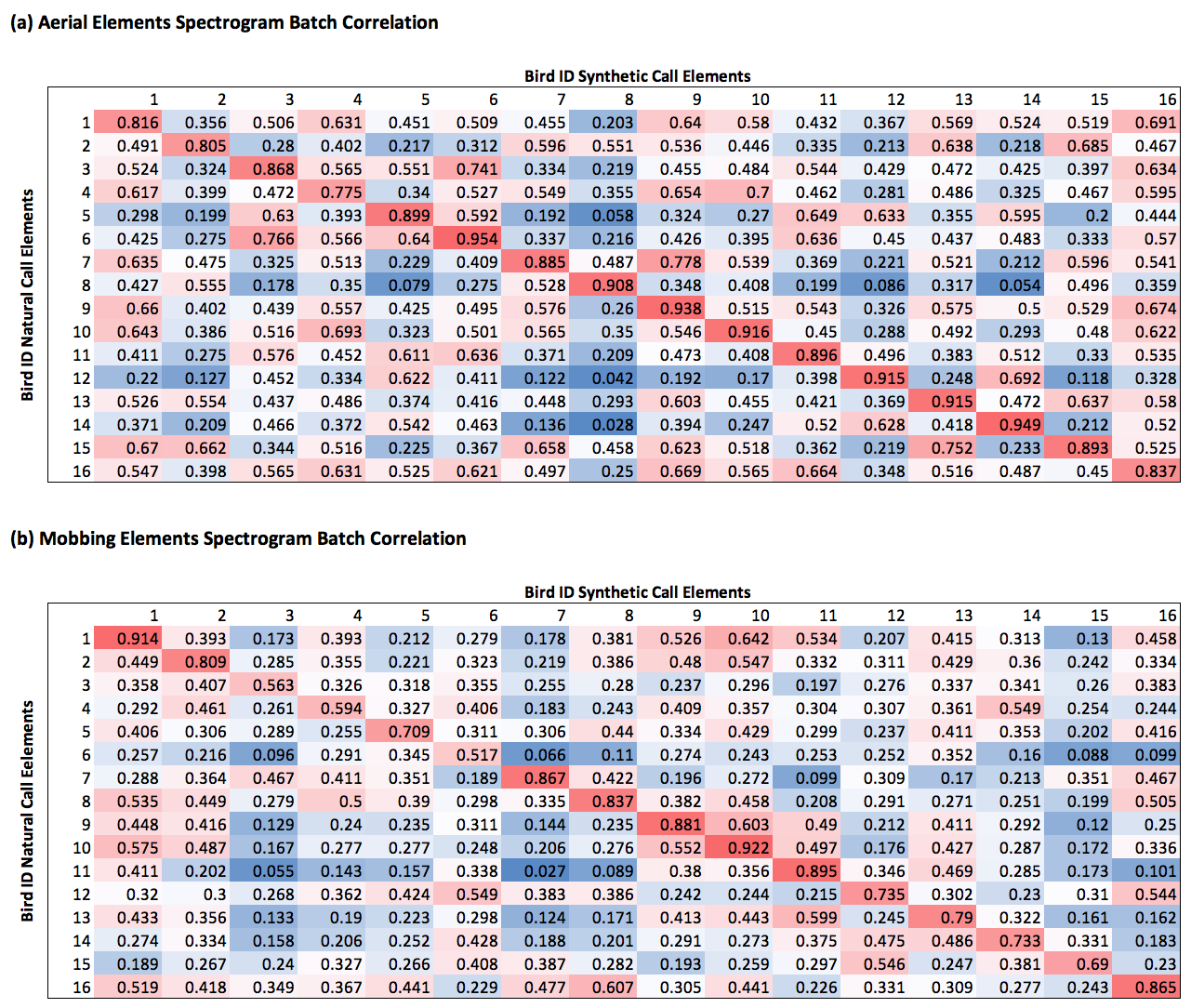


Table S3. PC1 and PC2 loadings from a principal component analysis examining acoustic features of aerial and mobbing calls. This analysis includes the full sample of recordings, with n = 32 aerial and 48 mob elements. Data are shown in Figure 3.

|  |  |  |
| --- | --- | --- |
| Acoustic Feature | PC1 | PC2 |
| Duration (ms) | 0.100 | -0.039 |
| Low Frequency (Hz) | 0.452 | -0.090 |
| High Frequency (Hz) | 0.205 | -0.676 |
| Peak Frequency (Hz) | 0.429 | -0.173 |
| Slope (Hz) | -0.397 | -0.147 |
| Frequency Irregularity (Hz) | -0.441 | -0.030 |
| Rate Frequency Modulation (cycles/sec) | -0.148 | -0.664 |
| Power Uniformity (dB) | 0.431 | 0.202 |

Table S4. Acoustic properties (mean ± SD, range) of aerial (n = 32) and mobbing call elements (n = 48). Results from linear mixed-effects models comparing aerial and mobbing calls, with and without the mobbing hook component included. Probability values are uncorrected; bold text indicates significance at 5% level after Bonferroni correction.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Measure | Aerial | Mobbing | Mobbing without hook a | Aerial vs mobbing | | Aerial vs mobbing without hook | |
| Low Frequency (kHz) | 8.68 ± 0.24  8.25-9.38 | 6.46 ± 0.52  5.25-7.69 | 7.26 ± 0.43  6.56-8.63 | *t*54= -21.92 | p **< 0.001** | *t*54=17.31 | p **< 0.001** |
| High Frequency (kHz) | 9.77 ± 0.32  9.19-10.69 | 9.31 ± 0.66  7.69-10.69 | 9.41 ± 0.68  7.69-10.88 | *t*54= -3.74 | p **< 0.001** | *t*54= 2.91 | p = 0.005 |
| Bandwidth (kHz) | 1.08 ± 0.21  0.75-1.69 | 2.85 ± 0.68  1.13-3.21 | 2.16 ± 0.48  1.13-3.19 | *t*54= 13.64 | p **< 0.001** | *t*54= 11.26 | p **< 0.001** |
| Peak Frequency (kHz) | 9.04 ± 0.36  8.44-9.94 | 7.42 ± 0.66  6.19-8.81 | 7.77 ± 0.52  6.75-9.38 | *t*54= -12.53 | p **< 0.001** | *t*54= 12.32 | p **< 0.001** |
| Slope (kHz) | -0.04 ± 0.34  -0.60-0.78 | -2.01 ± 0.84  -0.23- -3.88 | -2.01 ± 0.84  -0.23- -3.88 | *t*54= 12.07 | p **< 0.001** | *t*54= 12.07 | p **< 0.001** |
| Frequency  Irregularity (kHz) | 0.20 ± 0.08  0.09-0.41 | 0.84 ± 0.12  0.58-1.14 | 0.42 ± 0.13  0.14-0.70 | *t*54= 26.06 | p **< 0.001** | *t*54= -8.42 | p **< 0.001** |
| Rate of Frequency Modulation (Hz) | 97.0 ± 9.5  80.5-117.6 | 103.2 ± 8.9  86.4-123.7 | 103.6 ± 9.6  80.9-125.0 | *t*54= 6.50 | p **< 0.001** | *t*54= -5.55 | p **< 0.001** |
| Duration (ms) | 96.94 ± 26.66  52.00-165.00 | 91.65 ± 8.27  70.0-109.0 | 70.69 ± 9.72  47.0-90.0 | *t*54= -1.38 | p = 0.175 | *t*54= 6.26 | p **< 0.001** |
| Centre Time (ms) | 57.91 ± 19.09  34.00-104.00 | 50.88 ± 15.32  15.00-80.00 | 42.13 ± 11.63  15.0-66.0 | *t*54= -1.82 | p = 0.074 | *t*54= 4.60 | p **<** **0.001** |
| Time  Irregularity (ms) | 1.25 ± 0.37  0.71-2.38 | 1.29 ± 0.44  0.48-2.43 | 1.12 ± 0.47  0.45-2.56 | *t*54= 0.43 | p = 0.670 | *t*54= 1.47 | p = 0.148 |
| Aggregate Entropy (bits) | 2.8 ± 0.2  2.3-3.3 | 4.10 ± 0.37  3.08-4.67 | 3.69 ± 0.32  2.90-4.21 | *t*54= 16.91 | p **< 0.001** | *t*54= -12.75 | p **< 0.001** |
| Power Uniformity (dB) | 118.5 ± 1.2  116.0-121.8 | 111.7 ± 1.9  107.0-116.0 | 113.4 ± 2.65  106.9-119.4 | *t*54= -17.68 | p **< 0.001** | *t*54= 9.62 | p **< 0.001** |

a The slope is only calculated after the hook, so is unchanged when the hook is removed.