Supplementary Figures and Tables for article "The perceptual effects of manipulating nonlinear phenomena and source spectrum in synthetic nonverbal vocalizations"

Experiment 1

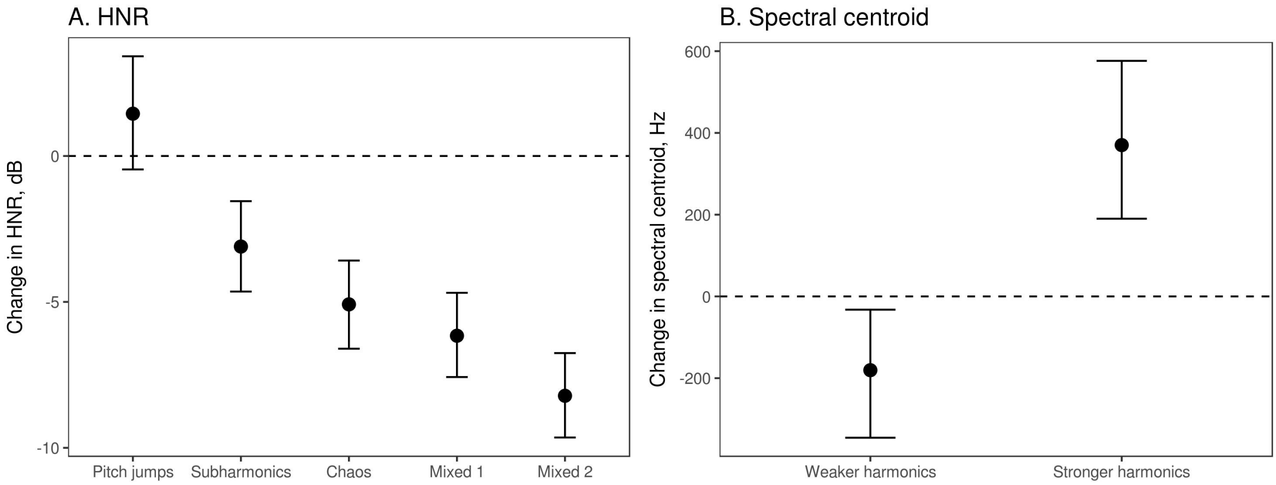

Fig. S1 Change in harmonics-to-noise ratio (HNR, panel A) and spectral centroid (panel B) after manipulating nonlinear effects (A) or rolloff of harmonics in the glottal source (B) compared to unmanipulated sound. Median of posterior distribution and 95% CI from mixed models with a random intercept per prototype sound. N = 28 prototypes with 144 sounds for (A) and 84 sounds for (B).

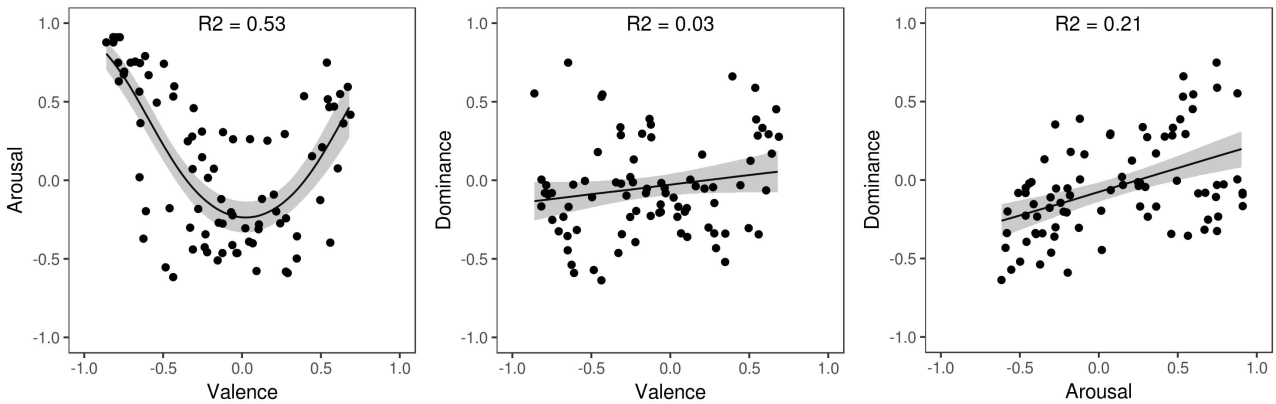

Fig. S2 Correlation between valence, arousal, and dominance ratings averaged per stimulus in Experiment 2: scatterplots and the best fit with 95% CI. Each point represents one stimulus. Valence-arousal: quadratic ($F(2,81)$ = 45.4, p < .001); valence-dominance: none ($F(2,81)$ = 3.0, p = .09); arousal-dominance: linear ($F(2,81)$ = 21.7, p < .001).

Table S1 The effect of adding nonlinear phenomena per call type

Scale	Call type	Predicted difference: median [95% CI]				
		Any vs. none	Pitch jumps vs. none	Subharmonics vs. none	Chaos vs. none	Mixed vs. none
Valence	**Cry**	**-0.09 [-0.16, -0.03]**	-0.02 [-0.15, 0.06]	-0.02 [-0.13, 0.04]	**-0.15 [-0.27, -0.03]**	**-0.13 [-0.21, -0.05]**
	Gasp	**-0.2 [-0.33, -0.08]**				**-0.2 [-0.33, -0.08]**
	Grunt	**-0.06 [-0.16, 0.06]**		-0.03 [-0.16, 0.11]	-0.02 [-0.18, 0.16]	-0.09 [-0.2, 0.05]
	Laugh	**-0.09 [-0.17, -0.02]**	-0.03 [-0.18, 0.09]	0 [-0.1, 0.11]	-0.1 [-0.22, 0.02]	**-0.16 [-0.27, -0.07]**
	Moan	**-0.18 [-0.31, -0.06]**		**-0.28 [-0.46, -0.07]**	0 [-0.17, 0.18]	**-0.23 [-0.37, -0.1]**
	Roar	**-0.12 [-0.22, -0.04]**	-0.01 [-0.14, 0.13]	-0.06 [-0.2, 0.05]	**-0.22 [-0.36, -0.09]**	**-0.17 [-0.28, -0.07]**
	Scream	**-0.19 [-0.3, -0.09]**	**-0.25 [-0.39, -0.1]**	-0.01 [-0.13, 0.11]	**-0.2 [-0.33, -0.08]**	**-0.25 [-0.37, -0.13]**
Arousal	**Cry**	0 [-0.03, 0.04]	0 [-0.1, 0.04]	0.02 [-0.02, 0.11]	0.01 [-0.02, 0.11]	0 [-0.05, 0.02]
	Gasp	0 [-0.06, 0.12]				0 [-0.06, 0.12]
	Grunt	0 [-0.07, 0.06]		0.03 [-0.04, 0.16]	0.01 [-0.09, 0.11]	-0.02 [-0.13, 0.03]
	Laugh	0 [-0.07, 0.07]	-0.01 [-0.13, 0.09]	0.07 [-0.02, 0.27]	0.02 [-0.06, 0.13]	-0.03 [-0.16, 0.02]
	Moan	0.02 [-0.04, 0.13]		0.03 [-0.03, 0.21]	0 [-0.14, 0.1]	0.02 [-0.04, 0.15]
	Roar	0 [-0.05, 0.06]	0 [-0.11, 0.11]	0.03 [-0.04, 0.15]	0.02 [-0.05, 0.14]	-0.02 [-0.12, 0.03]
	Scream	0 [-0.03, 0.05]	-0.01 [-0.12, 0.03]	0 [-0.08, 0.08]	0.01 [-0.03, 0.1]	0 [-0.04, 0.08]
Dominance	**Cry**	0.01 [-0.02, 0.05]	-0.01 [-0.11, 0.03]	0.02 [-0.02, 0.11]	0 [-0.05, 0.05]	0.02 [-0.01, 0.08]
	Gasp	0.07 [-0.02, 0.19]				0.07 [-0.02, 0.19]
	Grunt	0.04 [-0.01, 0.13]		0.03 [-0.06, 0.14]	0.01 [-0.06, 0.15]	0.05 [-0.01, 0.18]
	Laugh	0.01 [-0.05, 0.06]	-0.03 [-0.2, 0.05]	0.05 [-0.02, 0.19]	0 [-0.08, 0.08]	0.01 [-0.06, 0.08]
	Moan	0.05 [-0.02, 0.14]		0.04 [-0.03, 0.18]	-0.01 [-0.14, 0.06]	0.09 [-0.01, 0.21]
	Roar	0.02 [-0.04, 0.09]	-0.05 [-0.22, 0.03]	0.09 [-0.01, 0.26]	0.01 [-0.06, 0.15]	0.02 [-0.04, 0.09]
	Scream	0.01 [-0.04, 0.06]	-0.01 [-0.1, 0.1]	0 [-0.13, 0.09]	-0.01 [-0.11, 0.06]	0.04 [-0.02, 0.12]

Table S2 The effect of manipulating the strength of harmonics relative to F0 per call type

Scale	Call type	Predicted difference depending on the energy in harmonics: median [95% CI]		
		More vs. less	More vs. original	Original vs. less
Valence	Cry	-0.06 [-0.18, 0.02]	-0.04 [-0.16, 0.06]	-0.01 [-0.11, 0.06]
	Gasp	-0.09 [-0.27, 0.06]	-0.09 [-0.28, 0.1]	0 [-0.16, 0.16]
	Grunt	-0.06 [-0.23, 0.11]	-0.07 [-0.27, 0.11]	0.01 [-0.14, 0.2]
	Laugh	0.05 [-0.1, 0.27]	-0.02 [-0.19, 0.15]	0.07 [-0.06, 0.28]
	Moan	-0.12 [-0.33, 0.03]	-0.09 [-0.29, 0.1]	-0.03 [-0.22, 0.11]
	Roar	-0.04 [-0.18, 0.12]	0.05 [-0.1, 0.23]	-0.08 [-0.27, 0.04]
	Scream	-0.05 [-0.15, 0.04]	-0.03 [-0.14, 0.08]	-0.02 [-0.12, 0.07]
Arousal	Cry	**0.11 [0, 0.23]**	0.04 [-0.08, 0.16]	0.06 [-0.02, 0.19]
	Gasp	0.11 [-0.02, 0.26]	0.11 [-0.05, 0.3]	0.01 [-0.18, 0.14]
	Grunt	0.11 [-0.03, 0.25]	0.05 [-0.1, 0.21]	0.05 [-0.08, 0.2]
	Laugh	**0.13 [0, 0.3]**	0.05 [-0.1, 0.22]	0.08 [-0.04, 0.24]
	Moan	0.1 [-0.04, 0.24]	0.06 [-0.09, 0.23]	0.03 [-0.12, 0.17]
	Roar	0.07 [-0.1, 0.21]	-0.1 [-0.31, 0.08]	**0.17 [0.01, 0.38]**
	Scream	0.05 [-0.04, 0.14]	0.01 [-0.08, 0.11]	0.03 [-0.04, 0.13]
Dominance	Cry	0 [-0.03, 0.06]	0.01 [-0.03, 0.09]	0 [-0.07, 0.02]
	Gasp	0 [-0.08, 0.08]	0.01 [-0.07, 0.13]	-0.01 [-0.12, 0.05]
	Grunt	0.01 [-0.05, 0.12]	0.01 [-0.1, 0.12]	0 [-0.07, 0.12]
	Laugh	0.01 [-0.04, 0.15]	0.02 [-0.06, 0.17]	0 [-0.08, 0.08]
	Moan	0.01 [-0.04, 0.15]	0.02 [-0.05, 0.17]	0 [-0.1, 0.06]
	Roar	0 [-0.1, 0.07]	0 [-0.11, 0.1]	0 [-0.09, 0.07]
	Scream	0 [-0.12, 0.06]	0 [-0.11, 0.1]	0 [-0.1, 0.06]

Experiment 2

Table S3 The algorithm for transforming the marker's coordinates inside an equilateral triangle into relative weights of the vertices.
Taking as input Euclidean coordinates of each vertex and of the marked point, which is assumed (in this case forced) to lie within the triangle, do the following: 1. Calculate the Euclidean distance from the marker to each vertex D_vertex_{1-3}. 2. Calculate the Euclidean distance from the marker to each edge D_edge_{1-3}. If the marked point has coordinates (x0, y0) and the two relevant vertices have coordinates (x1, y1) and (x2, y2), the distance from the marker to the edge D_edge_i is given by: $D\_edge_i = \|(y2 - y1) * x0 - (x2 - x1) * y0 + x2 * y1 - y2 * x1\| / sqrt((x1 – x2) ^ 2 + (y1 – y2) ^ 2))$, where |...| is the absolute value (modulus). 3. Convert distances to edges into weights. For distances D_edge_{1-3}, $W\_edge\_raw_i = (sum(D\_edge_1, D\_edge_2, D\_edge_3) / (D\_edge_i + 0.0001)) ^ 2$ 4. Normalize weights $W\_edge\_raw_{1-3}$ to range from 0 to 100%: $W\_edge_i = W\_edge\_raw_i / (W\_edge_1 + W\_edge_2 + W\_edge_3) * 100$ 5. Find the edge closest to the marker and then adjust the distance to the vertex D_vertex_f which lies opposite to the closest edge: $D\_vertex\_adj_f = D\_vertex_f / sqrt((100 – W\_edge_c) / (100 – 33.3))$, where W_edge_c is the weight of the closest edge. Note that $D\_vertex\_adj_f$ approaches infinity as W_edge_c approaches 100, and therefore the weight of the vertex opposite to the closest edge approaches zero as the marker approaches the edge. The weights of edges are not allowed to become exactly zero due to the addition of a small constant in (3). If the marker is exactly equidistant to all three vertices, the weight of each vertex is ~33.3% and this adjustment has no effect. 6. Convert the adjusted distances $D\_edge\_adj_{1-3}$ into weights as in (3) and normalize to range from 0 to 100% as in (4).

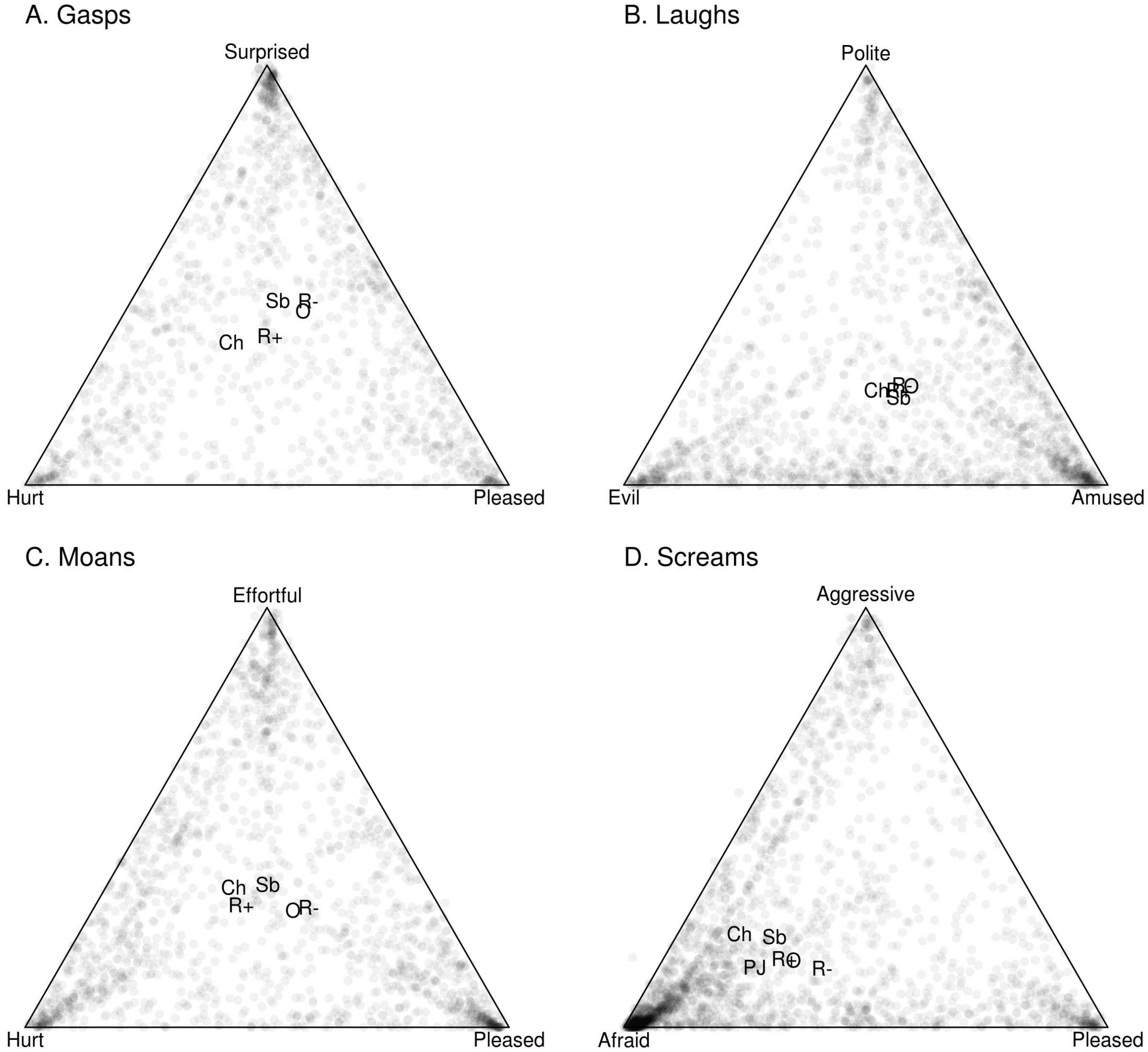

Fig. S3 Ratings on the triadic scale for each call type in standardized coordinates (the location of labels was random for each participant). The semi-transparent gray points show ratings from individual trials, while text labels inside the triangle show the center of gravity for each manipulation. O = original, R- = less energy in harmonics, R+ = more energy in harmonics, Sb = subharmonics, Ch = chaos, PJ = pitch jumps.

While the vertices attracted most clicks, there were also many ratings along the edges and closer to the middle of the triangle. Some combinations of emotions were more common than others: for example, for screams there were many responses between *Afraid* and *Aggresive* and between *Afraid* and *Pleased*, but few between *Pleased* and *Aggressive*. This is consistent with response patterns observed with similar sounds when each emotion had its own scale (Anikin & Persson, 2017). Qualitatively, the triadic scale was thus used appropriately and expressed combinations of emotions. However, the consistency of responses was low: comparing either the coordinates or label weights chosen for the same sound by different participants gives ICC = .17 for gasps, .14 for laughs, .21 for moans, and .25 for screams.

Table 4. Contrasts between the effect of different acoustic manipulations on the weight of emotion categories for each call type: median of posterior distribution (%) and 95% CI (repeated from the main text for ease of comparison with Table S5 below).

	Gasps				Laughs		
Contrast	**Hurt**	**Surprised**	**Pleased**	**Contrast**	**Evil**	**Polite**	**Amused**
Sb vs. O*	**7.1 [1.5, 12.7]**	0.5 [-7.4, 8.6]	**-7.6 [-13.1, -2.1]**	Sb vs. O	5.6 [-1, 12.3]	-1.9 [-6.7, 2.8]	-3.7 [-8.5, 1]
Ch vs. O	**23.3 [17.1, 29.7]**	**-9.5 [-18.3, -0.6]**	**-13.9 [-20.2, -7.8]**	Ch vs. O	**9.2 [2, 16.9]**	-1 [-6.2, 4.2]	**-8.3 [-13.6, -3]**
Ch vs. Sb	**16.2 [10.5, 22]**	**-10 [-17.9, -1.9]**	**-6.2 [-11.9, -0.7]**	Ch vs. Sb	3.7 [-2.9, 10.4]	0.9 [-3.7, 5.7]	**-4.7 [-9.4, 0]**
R+ vs. R-	**15 [8.7, 21.3]**	-7.2 [-16.4, 1.8]	**-7.8 [-14.2, -1.3]**	R+ vs. R-	1.8 [-5.3, 8.9]	-2.6 [-7.4, 2.6]	0.8 [-4.3, 5.7]
	Moans				**Screams**		
Contrast	**Hurt**	**Effortful**	**Pleased**	**Contrast**	**Afraid**	**Aggressive**	**Pleased**
Sb vs. O	3.1 [-1.9, 8]	**7.3 [2.3, 12.1]**	**-10.4 [-17.3, -3.5]**	Sb vs. O	2.1 [-1.6, 5.8]	**6.4 [2.7, 10.1]**	**-8.5 [-13.7, -3.3]**
Ch vs. O	**11.5 [6.4, 16.8]**	**5.7 [0.5, 10.9]**	**-17.3 [-24.7, -9.8]**	Ch vs. O	**9.7 [5.3, 13.8]**	**7 [2.8, 11.3]**	**-16.7 [-22.7, -10.8]**
Ch vs. Sb	**8.5 [3.2, 13.5]**	-1.5 [-6.6, 3.5]	-7 [-14.1, 0.4]	Ch vs. Sb	**7.6 [3.9, 11.3]**	0.6 [-3.1, 4.3]	**-8.2 [-13.3, -3]**
R+ vs. R-	**15.7 [10.7, 20.7]**	0.4 [-4.6, 5.5]	**-16.2 [-23.3, -9.1]**	R+ vs. R-	**8 [3.6, 12.3]**	2.4 [-1.9, 6.8]	**-10.4 [-16.4, -4.2]**
				PJ vs. O	**9 [4.4, 13.6]**	0.7 [-4, 5.3]	**-9.6 [-16.1, -3]**

* O = original, R- = less energy in harmonics, R+ = more energy in harmonics, Sb = subharmonics, Ch = chaos, PJ = pitch jumps

Table S4. Contrasts between the effect of different acoustic manipulations on the weight of emotion categories for each call type, controlling for harmonics-to-noise ratio (HNR) and spectral centroid: median of posterior distribution (%) and 95% CI. Cf. Table 50 above.

	Gasps				Laughs		
Contrast	**Hurt**	**Surprised**	**Pleased**	**Contrast**	**Evil**	**Polite**	**Amused**
Sb vs. O*	**7.1 [1, 12.9]**	-1.3 [-9.8, 7.2]	-5.8 [-11.8, 0.2]	Sb vs. O	1.7 [-5.6, 9]	-0.6 [-5.6, 4.8]	-1.2 [-6.5, 4.1]
Ch vs. O	**23.4 [16.5, 30.3]**	**-12 [-21.6, -2.4]**	**-11.4 [-18.5, -4.6]**	Ch vs. O	3.8 [-4.7, 12.7]**	1 [-5.2, 7.1]	-4.8 [-11, 1.4]
Ch vs. Sb	**16.4 [10.7, 22]**	**-10.8 [-18.7, -2.7]**	-5.6 [-11.5, 0.1]	Ch vs. Sb	2.1 [-4.8, 9]	1.5 [-3.4, 6.4]	-3.6 [-8.4, 1.2]
R+ vs. R-	**15.7 [8.1, 23.3]**	-4.4 [-15.3, 6.6]	**-11.3 [-19, -3.6]**	R+ vs. R-	-2 [-10.9, 7]	0.6 [-5.8, 6.8]	1.4 [-4.8, 7.7]
	Moans				**Screams**		
Contrast	**Hurt**	**Effortful**	**Pleased**	**Contrast**	**Afraid**	**Aggressive**	**Pleased**
Sb vs. O	1.2 [-4.3, 6.7]	**7.5 [2.1, 13]**	**-8.7 [-16.5, -1.1]**	Sb vs. O	**4.9 [0.1, 9.5]**	3.2 [-1.5, 7.8]	**-8 [-14.6, -1.5]**
Ch vs. O	**8.6 [2, 15.4]**	6.3 [-0.3, 13]	**-14.9 [-24.5, -5.5]**	Ch vs. O	**15.9 [8.2, 23.6]**	-0.2 [-7.6, 7.4]	**-15.8 [-25.8, -5.4]**
Ch vs. Sb	**7.4 [2.1, 12.8]**	-1.2 [-6.5, 4.1]	-6.2 [-13.8, 1.2]	Ch vs. Sb	**11 [5.9, 16.1]**	-3.3 [-8.3, 1.7]	**-7.7 [-14.5, -1]**
R+ vs. R-	**17.8 [11.4, 24.2]**	0.8 [-5.7, 7.2]	**-18.6 [-27.5, -9.6]**	R+ vs. R-	5.7 [-0.4, 11.5]	0.6 [-5.1, 6.3]	-6.2 [-14.4, 1.9]
				PJ vs. O	**8.9 [4.2, 13.4]**	1 [-3.7, 5.4]	**-9.8 [-16.2, -3.3]**

* O = original, R- = less energy in harmonics, R+ = more energy in harmonics, Sb = subharmonics, Ch = chaos, PJ = pitch jumps
** Where the models with and without controlling for HNR and spectral centroid disagree, in the sense that the magnitude of effect or the location of the 95% CI lead to substantively different interpretations, the cell is highlighted. For example, the model with HNR and spectral centroid predicts no effect of chaos on laughs, whereas the model without these variables predicts that chaos shifts the interpretation from "amused" to "evil".