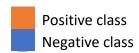
Zero-Shot Transfer Learning to Enhance Communication for Minimally Verbal Individuals with Autism using Naturalistic Data

Jaya Narain* & Kristina T. Johnson*, Rosalind Picard, Pattie Maes (*Equal Contribution)

AudioSet indices below match those in <u>class_labels_indices.csv</u> from the Features Dataset found here: <u>https://research.google.com/audioset/download.html</u>

As discussed in the text, the three model categories are Laughter, Negative Affect, and Self-Talk. Each positive class contains sounds that were similar (or as similar as possible) to the child's vocalizations for that category. Each negative class contains sounds that might confuse the model because they sounded similar to the child's vocalizations in a different category or appeared frequently in the dataset.

<u>Index</u>	<u>Label Name</u>	<u>Laugh</u>		<u>h</u>	Neg. Aff		Self-Talk			
0	Speech									
1	Male speech									
2	Female speech									
3	Child speech									
4	Conversation									
5	Narration									
6	Babbling									
7	Speech synthesizer									
8	Shout									
9	Bellow									
10	Whoop									
11	Yell									
12	Battle cry									
13	Children shouting									
14	Screaming									
15	Whispering									
16	Laughter									
17	Baby laughter									
18	Giggle									
19	Snicker									
20	Belly laugh									
21	Chuckle, chortle									
22	Crying, sobbing									
23	Infant cry									
24	Whimper									
25	Wail, moan									



26 Sigh 27 Singing 28 Choir 29 Yodeling 30 Chant 31 Mantra 32 Male singing 33 Female singing 34 Child singing 35 Synthetic singing 36 Rapping 37 Humming 38 Groan 39 Grunt 63 Clapping 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion							
28 Choir 29 Yodeling 30 Chant 31 Mantra 32 Male singing 33 Female singing 34 Child singing 35 Synthetic singing 36 Rapping 37 Humming 38 Groan 39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	26	Sigh					
29 Yodeling 30 Chant 31 Mantra 32 Male singing 33 Female singing 34 Child singing 35 Synthetic singing 36 Rapping 37 Humming 38 Groan 39 Grunt 63 Clapping 63 Clapping 69 70 Speech babble 137 Music 161 Percussion	27	Singing					
30 Chant 31 Mantra 32 Male singing 33 Female singing 34 Child singing 35 Synthetic singing 36 Rapping 37 Humming 38 Groan 39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	28	Choir					
31 Mantra 32 Male singing 33 Female singing 34 Child singing 35 Synthetic singing 36 Rapping 37 Humming 38 Groan 39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	29	Yodeling					
32 Male singing 33 Female singing 34 Child singing 35 Synthetic singing 36 Rapping 37 Humming 38 Groan 39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	30	Chant					
33 Female singing 34 Child singing 35 Synthetic singing 36 Rapping 37 Humming 38 Groan 39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	31	Mantra					
34 Child singing 35 Synthetic singing 36 Rapping 37 Humming 38 Groan 39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	32	Male singing					
35 Synthetic singing 36 Rapping 37 Humming 38 Groan 39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	33	Female singing					
36 Rapping 37 Humming 38 Groan 39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	34	Child singing					
37 Humming 38 Groan 39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	35	Synthetic singing					
38 Groan 39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	36	Rapping					
39 Grunt 45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	37	Humming					
## 15 ##	38	Groan					
45 Pant 63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	39	Grunt					
63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion							
63 Clapping 69 Crowd 70 Speech babble 137 Music 161 Percussion	45	Pant					
69 Crowd 70 Speech babble 137 Music 161 Percussion							
69 Crowd 70 Speech babble 137 Music 161 Percussion	63	Clapping					
70 Speech babble 137 Music 161 Percussion							
137 Music 161 Percussion	69	Crowd					
137 Music 161 Percussion	70	Speech babble					
161 Percussion							
161 Percussion	137	Music					
	161	Percussion					