

1 **The pitch of babies' cries predicts their voice pitch at age five**

2

3 Florence Levrero¹, Nicolas Mathevon¹, Katarzyna Pisanski², Erik Gustafsson^{1†} &

4 David Reby^{2*}

5

6 ¹ Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI CNRS UMR 9197,

7 University of Lyon/Saint-Etienne, France

8 ² Mammal Vocal Communication & Cognition Research Group, School of

9 Psychology, University of Sussex, United Kingdom

10 † Current address: Department of Psychology, University of Portsmouth, United

11 Kingdom

12

13

14

15 * Corresponding author: D. Reby (reby@sussex.ac.uk)

16

17

18

19 Running Head: Longitudinal study of voice pitch

20

21 Word count: 2533

22 **Abstract**

23 Voice pitch (fundamental frequency, F_0) is a key dimension of our voice that varies
24 between sexes after puberty, and also among individuals of the same sex both before
25 and after puberty. While a recent longitudinal study indicates that inter-individual
26 differences in voice pitch remain stable in men during adulthood and may even be
27 determined before puberty [1], whether these differences emerge in infancy remains
28 unknown. Here, using a longitudinal study design, we investigate the hypothesis that
29 inter-individual differences in F_0 are already present in the cries of pre-verbal babies.
30 While based on a small sample ($n = 15$), our results indicate that the F_0 of babies'
31 cries at 4 months of age may predict the F_0 of their speech utterances at 5 years of
32 age, explaining 41% of the inter-individual variance in voice pitch at that age in our
33 sample. We also found that the right-hand ratio of the length of their index to ring finger
34 (2D:4D digit ratio), which has been proposed to constitute an index of prenatal
35 testosterone exposure, was positively correlated with F_0 at both 4 months and 5 years
36 of age. These findings suggest that a substantial proportion of between-individual
37 differences in voice pitch, which convey important biosocial information about
38 speakers, may partly originate *in utero* and thus already be present soon after birth.

39

40 Keywords: voice pitch; fundamental frequency; baby cries; longitudinal study; gender

41 **Introduction**

42 While the fundamental frequency (F_0) of the adult human voice is highly sexually
43 dimorphic, there are no sex differences in the F_0 of babies' cries [2] or pre-pubertal
44 children's speech [3 for review]. In contrast, strong individual differences characterise
45 the pitch of adults' voices, leading to distinct F_0 distributions in men's and women's
46 speech (men's modal F_0 typically ranges between 80-175 Hz and women's between
47 160-270 Hz [3]) and in their nonverbal vocalisations (e.g. tennis grunts: men 230-400
48 Hz, women 370-910 [4]). Among pre-pubertal children, in the absence of sex
49 differences, F_0 distributions largely overlap and are characterised by wide ranges in
50 both preverbal cries (F_0 ranges 350-550 Hz in the cries of 3-month-old babies [2]) and
51 children's speech (215 to 320 Hz in 5 to 7-year-old's speech [5]).

52 Interestingly, a recent longitudinal study indicates that, at least in male speakers, inter-
53 individual differences in F_0 are stable after puberty and may even be determined
54 before puberty, with voice pitch at age 7 predicting up to 64% of the variance in pitch
55 in adulthood [1]. These observations suggest that inter-individual differences in F_0
56 arise early in life and are largely unaffected by puberty, and raise the possibility that
57 F_0 may even be determined before birth.

58 Here, we test whether the F_0 of preverbal babies' cries predicts the F_0 of their speech
59 as verbal pre-pubertal children. We predicted that the F_0 of babies' cries at 4 months
60 of age would correlate positively with the F_0 of their speech at 5 years of age. Because
61 lower 2D:4D ratios have been linked to higher fetal testosterone levels [6,7, but see 8]
62 and more masculine traits, we also predicted that 2D:4D ratios measured at age 5
63 would correlate positively with voice F_0 in both infancy and childhood.

64

Methods

65 *Participants*

66 We recorded the voices of 15 French children (6 females and 9 males) aged
67 between 4.2 and 5.9 years (mean 5.1 ± 0.7 years), for whom we already had
68 recordings of mild discomfort cries (bathing cries) obtained when they were between
69 2.7 and 5.1 months old (mean 3.8 ± 0.6 months) for the purpose of a previous
70 investigation of sex differences in babies' cries (involving 13 females and 15 males
71 [2]).

72 Participants were recorded in their own home, in the presence of one or both
73 parents. Parents were asked to measure the height and weight of their child prior to
74 the experimenters' visit. There were no significant sex differences in the age, height
75 and weight of children (all $p > 0.05$). Parental written consent and children's verbal
76 assent were systematically obtained.

77

78 *Audio Recordings and Analyses*

79 To obtain the F_0 of children's speech at 5 years of age, free speech was
80 recorded from child participants as they responded to simple, neutral questions about
81 themselves (i.e., their school, teachers, friends, family and hobbies). Children were
82 also recorded as they described an A4-sized, colour printed scene of the popular
83 cartoon "Peppa Pig" [9]. Recordings were performed using a Zoom H4n handheld
84 recorder positioned at approximately 30 cm from the participant. Sound files were
85 recorded at 44.1 kHz, 16 bits and saved in WAV format.

86 Acoustic editing and analysis were performed in PRAAT v.5.2.1 [10]. Recordings
87 were edited manually to remove all silences, nonverbal vocalisations (e.g., laughter,
88 loud breathing or nonverbal interjections) acute background noises and the
89 experimenter's questions, resulting in concatenated speech sequences ranging
90 between 19.9 and 73.3 seconds (mean 50.3 ± 14.1 s). The pulse-detection based
91 "voice report" function was used to extract the average F_0 of each sequence, with a
92 search range of 40-800 Hz.

93 Six crying sequences from each of these same children, taken from a database
94 of babies' cries used for a previous study (see [2] for detailed recording procedures)
95 were edited to remove short inspiration bouts between cries. The duration of the
96 sequences ranged between 5.2 and 9.8 seconds (mean 8.0 ± 1.1 s). Their average
97 F_0 was then measured using search ranges varying between 80-600 Hz and 80-1200
98 Hz, as informed by visual inspection of narrow band spectrograms (window length =
99 0.05 s). Individual F_0 means were computed by averaging F_0 measures across
100 sequences within individuals.

101

102 *Analysis of 2D:4D digit ratios*

103 Children were asked to position their hand with their palm up, flat against a 5 mm-
104 squared graph paper positioned on a solid surface (coffee or dinner table). A minimum
105 of two pictures of each hand were taken using an iPhone 5c smartphone camera, at
106 approximately 30 cm above the hand. The lengths of the index (2D) and ring (4D)
107 fingers were measured as their ventral length, i.e. the distance between the tip of the
108 finger to the mid-point on the ventral crease proximal to the palm by a researcher blind
109 to the subject's vocal parameters. Measures were performed on a computer screen

110 on two different photos for each hand (except for 3 hands where only one photo was
111 available), and the measured lengths were averaged for each finger. This enabled us
112 to calculate 2D:4D digit ratios for the left and right hands of each participant.

113

114 *Statistical Analyses*

115 Independent *t*-tests were performed to confirm the absence of sex differences in *F0*
116 (baby cry and child speech) and in 2D:4D ratios in our samples. A stepwise multiple
117 linear regression was then performed to investigate the relationships among: age,
118 body weight, body height, and baby cry *F0* as predictors, with child speech *F0* as the
119 outcome variable. Finally, two simple regressions were performed to investigate the
120 relationships between: 2D:4D ratios and cry *F0*, 2D:4D ratios and speech *F0*.

121

122 **Results**

123 *Descriptive statistics*

124 There were no significant sex differences in the mean *F0* of babies' cries, the mean
125 *F0* of children's speech, or their right and left hand 2D:4D ratios (table 1).

126

127 Table 1. Independent *t*-tests (two-sided) testing for sex differences in baby cry *F0* (Hz),
128 child speech *F0* (Hz), and child 2D:4D digit ratios. Range *F0* is reported for information.

	Sex (n)	Mean	(Range)	SE	<i>t</i> ¹	<i>df</i>	<i>p</i>
Baby cry <i>F0</i>	Male (9)	430.7	(383-533)	14.8	0.68	13	.511

	Female (6)	449.1	(387-543)	25.3			
Child speech F0	Male (9)	263.8	(216-312)	9.2	0.79	13	.445
	Female (6)	274.0	(254-305)	7.8			
Right 2D:4D	Male (9)	0.94	(0.87-1.0)	0.01	0.76	13	.461
	Female (6)	0.96	(0.90-1.02)	0.02			
Left 2D:4D	Male (9)	0.93	(0.89-1.02)	0.02	0.27	13	.788
	Female (6)	0.93	(0.88-0.98)	0.01			

129 1. Levene's Tests of Equality of Variances were nonsignificant, indicating that the
130 assumption of homogeneity of variance was not violated.

131

132 *Relationship between baby cry F0 and child speech F0*

133 A stepwise multiple linear regression showed that neither the age, weight nor height
134 at time of voice recording explained the mean F0 of children's speech. The only
135 predictor kept in the model was the mean F0 of their cries as babies indicating that
136 inter-individual differences in the mean F0 of babies' cries explained 40.6% of the
137 inter-individual differences in the F0 of their speech as children ($r = .637$, $p = 0.011$;
138 Table 2, Figure 1).

139 Table 2. Stepwise multiple linear regression examining the effects of age and body
140 size at five-year old, as well as F0 as a baby, on the mean F0 of children's speech.

	B	SE	t	p
Constant	134.332	45.083	2.980	.011
Baby F0	0.305	0.102	2.980	.011

Excluded variables:	Beta in	t	p
Age	-0.31	1.497	.160
Weight	-0.043	0.190	.852
Height	-0.081	0.353	.730

141 Model summary $F(1,13)=8.9$, $p=0.011$, $R^2 = 0.406$. The assumptions of normality and
 142 homoscedasticity are met.

143

144 *Relationships between 2D:4D ratio and F0*

145 Simple regressions showed that the right hand 2D:4D digit ratio in children was a
 146 significant predictor of the mean $F0$ of their speech ($r=.648$, $p=0.009$; Figure 2). Inter-
 147 individual differences in children's right 2D:4D ratios explained 42.0% of inter-
 148 individual differences in the $F0$ of their speech as children, and 26.8% of inter-
 149 individual differences in the $F0$ of their baby cries. In contrast, left hand 2D:4D ratios
 150 were not significant predictors of $F0$ at either age (both $p>0.05$). See Electronic
 151 Supplementary Materials (Table S1) for full descriptive correlation matrix.

152

153 **Discussion**

154 Pitch is a highly salient feature of the human voice that affects listeners' perceptions
 155 of femininity and masculinity in babies' cries [2], children's speech [11] and adult
 156 speech [12,13]. Here, we found that the pitch of babies' cries at 4 months of age was
 157 a significant and substantial predictor of the pitch of their speech at 5 years of age.

158 This indicates that inter-individual differences in early, preverbal infant voice pitch may
159 remain partly preserved in later pre-pubertal speech, likely reflecting inter-individual
160 differences in vocal fold length (the key factor affecting voice pitch [14]). As recent
161 work indicates that inter-individual differences in F_0 at age 7 also extend into adulthood
162 in males [1], the present results suggest that adult inter-individual differences in voice
163 pitch may - at least partly - arise very early in life. In our sample, there were no sex
164 differences in the F_0 of babies' cries and children's speech, corroborating previous
165 work showing that voice pitch is not sexually dimorphic before puberty [2,5], and
166 justifying the validity of pooling data across sexes in our analyses.

167 We also found that right hand 2D:4D digit ratios at age 5 correlated positively with
168 voice pitch in both infancy and childhood. The fact that correlations between digit ratios
169 and voice pitch were only significant with the right hand is consistent with previous
170 observations that 2D:4D ratio is more sexually dimorphic and a better predictor of
171 prenatal androgens [6] and other indexical traits (e.g. sperm count [15] or athletic
172 ability [16]) in the right than left hand. Thus, inter-individual differences in voice F_0 and
173 right hand 2D:4D digit ratio may share a common origin, potentially reflecting exposure
174 to prenatal hormonal levels, and as such index a cluster of attributes related to gender,
175 masculinity and dominance throughout the lifespan. In our sample, 2D:4D ratios did
176 not differ between boys and girls, also consistent with previous reports that sex
177 differences in 2D:4D ratios in pre-pubertal children are typically very small (0.01 in
178 children aged two-years [6] and 6-14 years [17]) and often nonsignificant [6].

179 Given the small sample size in our investigation, constrained by the difficulty of
180 obtaining longitudinal data on the voice of human participants, further investigations
181 involving a larger sample, and in particular more female participants, are now clearly

182 required to confirm these correlations within both sexes. Like voice F_0 , 2D:4D digit
183 ratios are sexually dimorphic in adults [17], and lower 2D:4D ratios have been shown
184 to predict greater facial masculinity among men [18]. Future work could investigate if
185 2D:4D digit ratios are also predictors of voice pitch in adult speakers.

186

187

188 **Figure Captions**

189 **Figure 1.** Relationships between the mean F_0 of babies' cries around 4 months of age
190 and the mean F_0 of their speech around the age of 5 years. Boys ($n = 9$) are
191 represented as black dots and girls ($n = 6$) as grey dots. Dotted lines represent
192 trendlines for each sex. Statistics are reported for the pooled-sex trendline (solid
193 black).

194 **Figure 2.** Relationships between children's right hand 2D:4D ratio at around 5 years
195 of age and: (1) the mean F_0 of their cries as babies (around 4 months old; circles),
196 and (2) the mean F_0 of their speech as children (around 5 years old; squares). Boys
197 ($n = 9$) are represented as black markers and girls ($n = 6$) as grey markers. Dotted
198 lines represent trendlines for each sex. Statistics are reported for the pooled-sex
199 trendlines (solid black).

200

201 **Author contributions**

202 DR, FL & NM designed the investigation. FL, DR & EG collected the data. DR
203 performed the acoustic analyses. DR & KP performed the statistical analyses. DR and

204 FL performed the digit ratio analyses. DR, FL, KP & NM wrote the manuscript. The
205 manuscript was reviewed and approved by all authors, who agree to be accountable
206 for the work.

207 **Competing Interests**

208 The authors report no competing interests.

209 **Data accessibility**

210 The datasets supporting this article are available as Electronic Supplementary Material
211 (ESM1).

212

213 **Acknowledgements**

214 We thank all the children and their parents who agreed to take part in our study.

215

216 **Funding statement**

217 This work was partly supported by a Leverhulme Trust grant to DR (RPG-2016-396),
218 the University Lyon Saint-Etienne (research sabbatical to FL) and a Institut
219 Universitaire de France grant to NM. DR was also supported by a visiting professorship
220 from University Lyon Saint-Etienne during part of the project. KP was supported by
221 funding from the European Union's Horizon 2020 research and innovation programme
222 under the Marie Skłodowska-Curie grant agreement No 655859.

223 **Ethics**

224 The study was reviewed and approved by the Sciences and Technology Cross-
225 Schools Research Ethics Committee (C-REC) of the University of Sussex (ER-REBY-
226 2)

227 **References**

- 228 1. Fouquet M, Pisanski K, Mathevon N, Reby D. 2016 Seven and up: individual
229 differences in male voice fundamental frequency emerge before puberty and
230 remain stable throughout adulthood. *R. Soc. Open Sci.* **3**, 160395.
- 231 2. Reby D, Levréro F, Gustafsson E, Mathevon N. 2016 Sex stereotypes influence
232 adults' perception of babies' cries. *BMC Psychol.* **4**, 19.
- 233 3. Kreiman J, Sidtis D. 2011 *Foundations of voice studies: An interdisciplinary*
234 *approach to voice production and perception*. Wiley-Blackwell.
- 235 4. Raine J, Pisanski K, Reby D. 2017 Tennis grunts communicate acoustic cues to
236 sex and contest outcome. *Anim. Behav.* **130**, 47–55.
- 237 5. Lee S, Potamianos A, Narayanan S. 1999 Acoustics of children's speech:
238 Developmental changes of temporal and spectral parameters. *J. Acoust. Soc.*
239 *Am.* **105**, 1455–1468.
- 240 6. Lutchmaya S, Baron-Cohen S, Raggatt P, Knickmeyer R, Manning JT. 2004 2nd
241 to 4th digit ratios, fetal testosterone and estradiol. *Early Hum. Dev.* **77**, 23–28.
- 242 7. Zheng Z, Cohn MJ. 2011 Developmental basis of sexually dimorphic digit ratios.
243 *Proc. Natl. Acad. Sci.* **108**, 16289–16294.
- 244 8. Berenbaum SA, Bryk KK, Nowak N, Quigley CA, Moffat S. 2009 Fingers as a
245 marker of prenatal androgen exposure. *Endocrinology* **150**, 5119–5124.

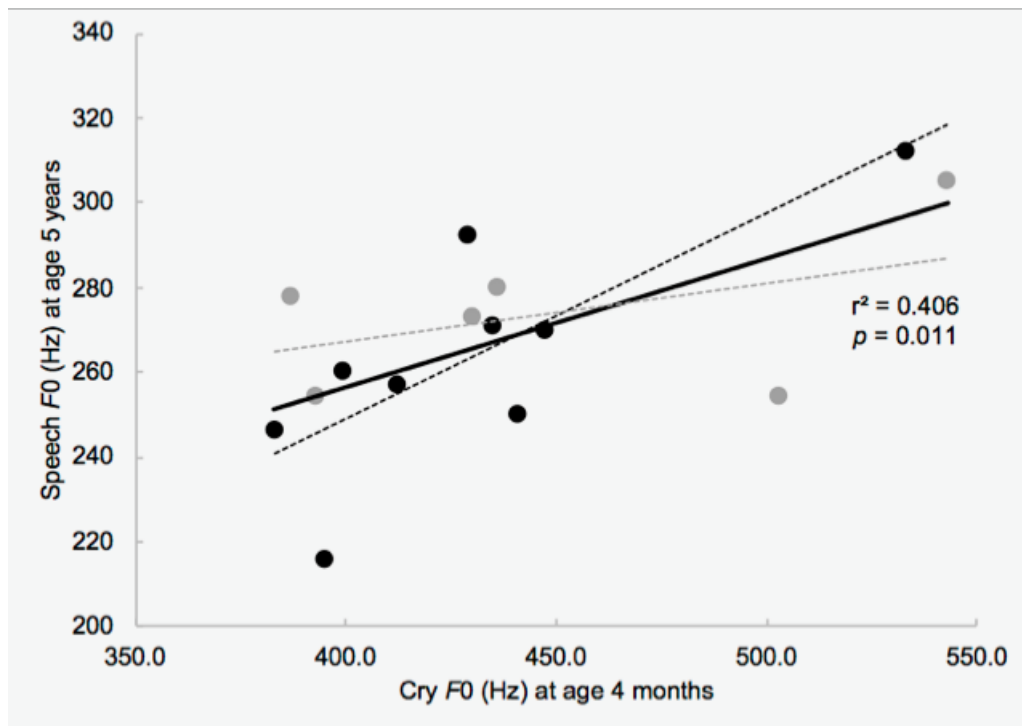
- 246 9. Astley N, Baker M. 2004 Peppa Pig [television series]
- 247 10. Boersma P, Weenink D. 2016 *Praat: Doing phonetics by computer v 6.0.21*. See
248 www.praat.org.
- 249 11. Cartei V, Reby D. 2013 Effect of formant frequency spacing on perceived gender
250 in pre-pubertal children's voices. *PLoS ONE* **8**, e81022.
- 251 12. Pisanski K, Rendall D. 2011 The prioritization of voice fundamental frequency or
252 formants in listeners' assessments of speaker size, masculinity, and
253 attractiveness. *J. Acoust. Soc. Am.* **129**, 2201.
- 254 13. Cartei V, Bond R, Reby D. 2014 What makes a voice masculine: Physiological
255 and acoustical correlates of women's ratings of men's vocal masculinity. *Horm.*
256 *Behav.* **66**, 569–576.
- 257 14. Titze IR. 2011 Vocal fold mass is not a useful quantity for describing f₀ in
258 vocalization. *J. Speech Lang. Hear. Res.* **54**, 520–522.
- 259 15. Manning JT, Scutt D, Wilson J, Lewis-Jones DI. 1998 The ratio of 2nd to 4th digit
260 length: a predictor of sperm numbers and concentrations of testosterone,
261 luteinizing hormone and oestrogen. *Hum. Reprod. Oxf. Engl.* **13**, 3000–3004.
- 262 16. Manning JT, Taylor RP. 2001 Second to fourth digit ratio and male ability in sport:
263 implications for sexual selection in humans. *Evol. Hum. Behav.* **22**, 61–69.
- 264 17. Manning JT, Stewart A, Bundred PE, Trivers RL. 2004 Sex and ethnic
265 differences in 2nd to 4th digit ratio of children. *Early Hum. Dev.* **80**, 161–168.

266 18. Meindl K, Windhager S, Wallner B, Schaefer K. 2012 Second-to-fourth digit ratio
267 and facial shape in boys: the lower the digit ratio, the more robust the face. *Proc.*
268 *R. Soc. Lond. B Biol. Sci.*, rspb20112351.

269

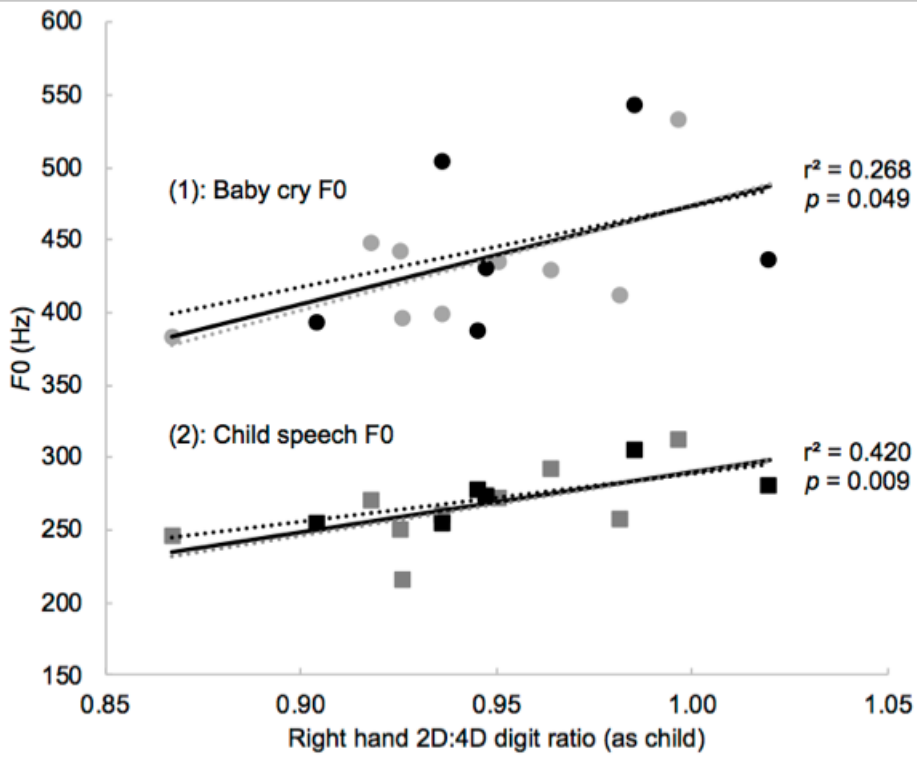
270 **Figure 1**

271



272
273
274
275
276

277 **Figure 2**



278
279