BIOLOGY LETTERS

rsbl.royalsocietypublishing.org

Research

Cite this article: Levrero F, Mathevon N, Pisanski K, Gustafsson E, Reby D. 2018 The pitch of babies' cries predicts their voice pitch at age 5. *Biol. Lett.* **14**: 20180065. http://dx.doi.org/10.1098/rsbl.2018.0065

Received: 29 January 2018
Accepted: 12 June 2018

Subject Areas:
behaviour

Keywords:
voice pitch, fundamental frequency, baby cries, longitudinal study, gender

Author for correspondence:
David Reby
e-mail: reby@sussex.ac.uk

†Present address: Department of Psychology, University of Portsmouth, Portsmouth PO1 2DY, UK.

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4144289.

THE ROYAL SOCIETY PUBLISHING

Animal behaviour

The pitch of babies' cries predicts their voice pitch at age 5

Florence Levrero[1], Nicolas Mathevon[1], Katarzyna Pisanski[2,3], Erik Gustafsson[1,†] and David Reby[2]

[1]Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI CNRS UMR 9197, University of Lyon/Saint-Etienne, Saint-Etienne, 42100, France
[2]Mammal Vocal Communication & Cognition Research Group, School of Psychology, University of Sussex, Brighton BN1 9QH, UK
[3]Institute of Psychology, University of Wrocław, 50-527 Wrocław, Poland

NM, 0000-0003-0219-6601; KP, 0000-0003-0992-2477; DR, 0000-0001-9261-1711

Voice pitch (fundamental frequency, F_0) is a key dimension of our voice that varies between sexes after puberty, and also among individuals of the same sex both before and after puberty. While a recent longitudinal study indicates that inter-individual differences in voice pitch remain stable in men during adulthood and may even be determined before puberty (Fouquet *et al.* 2016 *R. Soc. open sci.* **3**, 160395. (doi:10.1098/rsos.160395)), whether these differences emerge in infancy remains unknown. Here, using a longitudinal study design, we investigate the hypothesis that inter-individual differences in F_0 are already present in the cries of pre-verbal babies. While based on a small sample ($n = 15$), our results indicate that the F_0 of babies' cries at 4 months of age may predict the F_0 of their speech utterances at 5 years of age, explaining 41% of the inter-individual variance in voice pitch at that age in our sample. We also found that the right-hand ratio of the length of their index to ring finger (2D : 4D digit ratio), which has been proposed to constitute an index of prenatal testosterone exposure, was positively correlated with F_0 at both 4 months and 5 years of age. These findings suggest that a substantial proportion of between-individual differences in voice pitch, which convey important biosocial information about speakers, may partly originate *in utero* and thus already be present soon after birth.

1. Introduction

While the fundamental frequency (F_0) of the adult human voice is highly sexually dimorphic, there are no sex differences in the F_0 of babies' cries [1] or pre-pubertal children's speech ([2] for review). By contrast, strong individual differences characterize the pitch of adults' voices, leading to distinct F_0 distributions in men's and women's speech (men's modal F_0 typically ranges between 80 and 175 Hz and women's between 160 and 270 Hz [2]) and in their non-verbal vocalizations (e.g. tennis grunts: men 230–400 Hz, women 370–910 [3]). Among pre-pubertal children, in the absence of sex differences, F_0 distributions largely overlap and are characterized by wide ranges in both pre-verbal cries (F_0 ranges 350–550 Hz in the cries of 3-month-old babies [1]) and children's speech (215 to 320 Hz in 5- to 7-year-olds' speech [4]).

Interestingly, a recent longitudinal study indicates that, at least in male speakers, inter-individual differences in F_0 are stable after puberty and may even be determined before puberty, with voice pitch at age 7 predicting up to 64% of the variance in pitch in adulthood [5]. These observations suggest that inter-individual differences in F_0 arise early in life and are largely unaffected by puberty, and raise the possibility that F_0 may even be determined before birth.

Table 1. Independent *t*-tests (two-sided) testing for sex differences in baby cry F_0 (Hz), child speech F_0 (Hz), and child 2D : 4D digit ratios. Ranges in F_0 and 2D : 4D ratios are reported for information.

	sex (*n*)	mean	range	s.e.	T^a	d.f.	*p*-value
baby cry F_0	male (9)	430.7	383–533	14.8	0.68	13	0.511
	female (6)	449.1	387–543	25.3			
child speech F_0	male (9)	263.8	216–312	9.2	0.79	13	0.445
	female (6)	274.0	254–305	7.8			
right 2D : 4D	male (9)	0.94	0.87–1.0	0.01	0.76	13	0.461
	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	0.788
	female (6)	0.93	0.88–0.98	0.01			

[a]Levene's tests of equality of variances were non-significant, indicating that the assumption of homogeneity of variance was not violated.

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

2. Material and methods

(a) Participants

We recorded the voices of 15 French children (six females and nine males) aged between 4.2 and 5.9 years (mean 5.1 ± 0.7 years), for whom we already had recordings of mild discomfort cries (bathing cries) obtained when they were between 2.7 and 5.1 months old (mean 3.8 ± 0.6 months) for the purpose of a previous investigation of sex differences in babies' cries (involving 13 females and 15 males [1]).

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

(b) Audio recordings and analyses

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

Acoustic editing and analysis were performed in Praat v. 6.0.21 [10]. Recordings were edited manually to remove all silences, non-verbal vocalizations (e.g. laughter, loud breathing or non-verbal interjections) acute background noises and the experimenter's questions, resulting in concatenated speech sequences ranging between 19.9 and 73.3 s (mean 50.3 ± 14.1 s). The pulse-detection-based 'voice report' function was used to extract the average F_0 of each sequence, with a search range of 40–800 Hz.

Six crying sequences from each of these same children, taken from a database of babies' cries used for a previous study (see [1] for detailed recording procedures) were edited to remove short inspiration bouts between cries. The duration of the sequences ranged between 5.2 and 9.8 s (mean 8.0 ± 1.1 s). Their average F_0 was then measured using search ranges varying between 80–600 Hz and 80–1200 Hz, as informed by visual inspection of narrow band spectrograms (window length = 0.05 s). Individual F_0 means were computed by averaging F_0 measures across sequences within individuals.

(c) Analysis of 2D : 4D digit ratios

Children were asked to position their hand with their palm up, flat against a 5 mm-squared graph paper positioned on a solid surface (coffee or dinner table). A minimum of two pictures of each hand were taken using an iPhone 5c smartphone camera, at approximately 30 cm above the hand. The lengths of the index (2D) and ring (4D) fingers were measured as their ventral length, i.e. the distance between the tip of the finger to the mid-point on the ventral crease proximal to the palm by a researcher blind to the subject's vocal parameters. Measures were performed on a computer screen on two different photos for each hand (except for three hands where only one photo was available), and the measured lengths were averaged for each finger. This enabled us to calculate 2D : 4D digit ratios for the left and right hands of each participant.

(d) Statistical analyses

Independent *t*-tests were performed to confirm the absence of sex differences in F_0 (baby cry and child speech) and in 2D : 4D ratios in our samples. A stepwise multiple linear regression was then performed to investigate the relationships among age, body weight, body height and baby cry F_0 as predictors, with child speech F_0 as the outcome variable. Finally, two simple regressions were performed to investigate the relationships between 2D : 4D ratios and cry F_0, 2D : 4D ratios and speech F_0.

3. Results

(a) Descriptive statistics

There were no significant sex differences in the mean F_0 of babies' cries, the mean F_0 of children's speech, or their right- and left-hand 2D : 4D ratios (table 1).

(b) Relationship between baby cry F_0 and child speech F_0

A stepwise multiple linear regression showed that neither the age, weight nor height at the time of voice recording

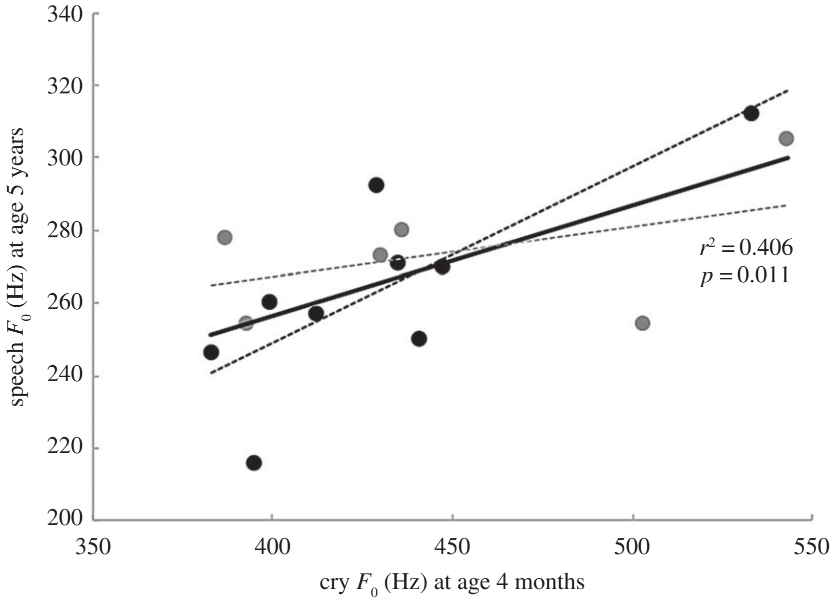

Figure 1. Relationships between the mean F_0 of babies' cries around 4 months of age and the mean F_0 of their speech around the age of five years. Boys ($n = 9$) are represented as black dots and girls ($n = 6$) as grey dots. Dotted lines represent trendlines for each sex. Statistics are reported for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

Table 2. Stepwise multiple linear regression examining the effects of age and body size at age at 5, as well as F_0 as a baby, on the mean F_0 of children's speech.

	β	s.e.	*t*	*p*-value
constant	134.332	45.083	2.980	0.011
baby F_0	0.305	0.102	2.980	0.011
excluded variables	**β in**		***t***	***p*-value**
age	−0.31		1.497	0.160
weight	−0.043		0.190	0.852
height	−0.081		0.353	0.730

Model summary: $F_{1,13} = 8.9$, $p = 0.011$, $R^2 = 0.406$. The assumptions of normality and homoscedasticity were met.

explained the mean F_0 of children's speech. The only predictor kept in the model was the mean F_0 of their cries as babies, indicating that inter-individual differences in the mean F_0 of babies' cries explained 40.6% of the inter-individual differences in the F_0 of their speech as children ($r = 0.637$, $p = 0.011$; table 2 and figure 1).

(c) Relationships between 2D : 4D ratio and F_0

Simple regressions showed that the right-hand 2D : 4D digit ratio in children was a significant predictor of the mean F_0 of their speech ($r = 0.648$, $p = 0.009$; figure 2). Inter-individual differences in children's right-hand 2D : 4D ratios explained 42.0% of inter-individual differences in the F_0 of their speech as children, and 26.8% of inter-individual differences in the F_0 of their baby cries. By contrast, left-hand 2D : 4D ratios were not significant predictors of F_0 at either age (both $p > 0.05$). See electronic supplementary material, table S1 for full descriptive correlation matrix.

4. Discussion

Pitch is a highly salient feature of the human voice that affects listeners' perceptions of femininity and masculinity in babies' cries [1], children's speech [11] and adult speech [12,13]. Here, we found that the pitch of babies' cries at 4 months of age was a significant and substantial predictor of the pitch of their speech at 5 years of age. This indicates that inter-individual differences in early, pre-verbal infant voice pitch may remain partly preserved in later pre-pubertal speech, likely reflecting inter-individual differences in vocal fold length (the key factor affecting voice pitch [14]). As recent work indicates that inter-individual differences in F_0 at age 7 also extend into adulthood in males [5], the present results suggest that adult inter-individual differences in voice pitch may—at least partly—arise very early in life. In our sample, there were no sex differences in the F_0 of babies' cries and children's speech, corroborating previous work showing that voice pitch is not sexually dimorphic before puberty [1,4], and justifying the validity of pooling data across sexes in our analyses.

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

Given the small sample size in our investigation, constrained by the difficulty of obtaining longitudinal data on

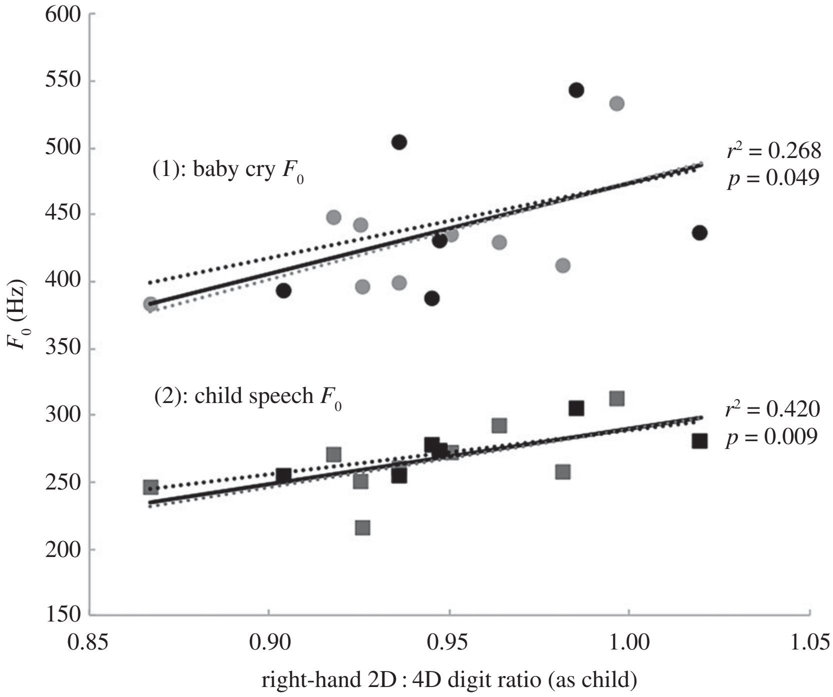

Figure 2. Relationships between children's right-hand 2D : 4D ratio at around 5 years of age and: (1) the mean F_0 of their cries as babies (around 4 months old; circles) and (2) the mean F_0 of their speech as children (around 5 years old; squares). Boys ($n = 9$) are represented as black markers and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

the voice of human participants, further investigations involving a larger sample, and in particular more female participants, are now clearly required to confirm these correlations within both sexes. Like voice F_0, 2D : 4D digit ratios are sexually dimorphic in adults [17], and lower 2D : 4D ratios have been shown to predict greater facial masculinity among men [18]. Future work could investigate whether 2D : 4D digit ratios are also predictors of voice pitch in adult speakers.

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

Data accessibility. The datasets supporting this article are available as electronic supplementary material, S1).

Authors' contributions. D.R., F.L. and N.M. designed the investigation. F.L., D.R. and E.G. collected the data. D.R. performed the acoustic analyses. D.R. and K.P. performed the statistical analyses. D.R. and F.L. performed the digit ratio analyses. D.R., F.L., K.P. and N.M. wrote the manuscript. The manuscript was reviewed and approved by all authors, who agreed to be accountable for the work.

Competing interests. We declare we have no competing interests.

Funding. This work was partly supported by a Leverhulme Trust grant to D.R. (RPG-2016-396), the University Lyon Saint-Etienne (research sabbatical to F.L.) and a Institut Universitaire de France grant to N.M. D.R. was also supported by a visiting professorship from University Lyon Saint-Etienne during part of the project. K.P. was supported by funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 655859.

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

References

1. Reby D, Levréro F, Gustafsson E, Mathevon N. 2016 Sex stereotypes influence adults' perception of babies' cries. *BMC Psychol.* **4**, 19. (doi:10.1186/s40359-016-0123-6)
2. Kreiman J, Sidtis D. 2011 *Foundations of voice studies: an interdisciplinary approach to voice production and perception*. Hoboken, NJ: Wiley-Blackwell.
3. Raine J, Pisanski K, Reby D. 2017 Tennis grunts communicate acoustic cues to sex and contest outcome. *Anim. Behav.* **130**, 47–55. (doi:10.1016/j.anbehav.2017.06.022)
4. Lee S, Potamianos A, Narayanan S. 1999 Acoustics of children's speech: developmental changes of temporal and spectral parameters. *J. Acoust. Soc. Am.* **105**, 1455–1468. (doi:10.1121/1.426686)
5. Fouquet M, Pisanski K, Mathevon N, Reby D. 2016 Seven and up: individual differences in male voice fundamental frequency emerge before puberty and remain stable throughout adulthood. *R. Soc. open sci.* **3**, 160395. (doi:10.1098/rsos.160395)
6. Lutchmaya S, Baron-Cohen S, Raggatt P, Knickmeyer R, Manning JT. 2004 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Hum. Dev.* **77**, 23–28. (doi:10.1016/j.earlhumdev.2003.12.002)
7. Zheng Z, Cohn MJ. 2011 Developmental basis of sexually dimorphic digit ratios. *Proc. Natl Acad. Sci. USA* **108**, 16289–16294. (doi:10.1073/pnas.1108312108)
8. Berenbaum SA, Bryk KK, Nowak N, Quigley CA, Moffat S. 2009 Fingers as a marker of prenatal androgen exposure. *Endocrinology* **150**, 5119–5124. (doi:10.1210/en.2009-0774)
9. Astley N, Baker M. 2004 Peppa Pig [television series].
10. Boersma P, Weenink D. 2016 *Praat: Doing phonetics by computer v 6.0.21*. See www.praat.org.
11. Cartei V, Reby D. 2013 Effect of formant frequency spacing on perceived gender in pre-pubertal children's voices. *PLoS ONE* **8**, e81022. (doi:10.1371/journal.pone.0081022)
12. Pisanski K, Rendall D. 2011 The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *J. Acoust. Soc. Am.* **129**, 2201. (doi:10.1121/1.3552866)

13. Cartei V, Bond R, Reby D. 2014 What makes a voice masculine: physiological and acoustical correlates of women's ratings of men's vocal masculinity. *Horm. Behav.* **66**, 569–576. (doi:10.1016/j.yhbeh.2014.08.006)
14. Titze IR. 2011 Vocal fold mass is not a useful quantity for describing F_0 in vocalization. *J. Speech Lang. Hear. Res.* **54**, 520–522. (doi:10.1044/1092-4388(2010/09-0284))
15. Manning JT, Scutt D, Wilson J, Lewis-Jones DI. 1998 The ratio of 2nd to 4th digit length: a predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen. *Hum. Reprod. Oxf. Engl.* **13**, 3000–3004. (doi:10.1093/humrep/13.11.3000)
16. Manning JT, Taylor RP. 2001 Second to fourth digit ratio and male ability in sport: implications for sexual selection in humans. *Evol. Hum. Behav.* **22**, 61–69. (doi:10.1016/S1090-5138(00)00063-5)
17. Manning JT, Stewart A, Bundred PE, Trivers RL. 2004 Sex and ethnic differences in 2nd to 4th digit ratio of children. *Early Hum. Dev.* **80**, 161–168. (doi:10.1016/j.earlhumdev.2004.06.004)
18. Meindl K, Windhager S, Wallner B, Schaefer K. 2012 Second-to-fourth digit ratio and facial shape in boys: the lower the digit ratio, the more robust the face. *Proc. R. Soc. B* **279**, 2457–2463. (doi:10.1098/rspb.2011.2351)