Sussex Research Online

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

Article (Accepted Version)

Levrero, Florence, Mathevon, Nicholas, Pisanski, Katarzyna, Gustafsson, Erik and Reby, David (2018) The pitch of babies' cries predicts their voice pitch at age five. Biology Letters, 14 (7). ISSN 1744-9561

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/76808/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk

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

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

[1] Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI CNRS UMR 9197, University of Lyon/Saint-Etienne, France

[2] Mammal Vocal Communication & Cognition Research Group, School of Psychology, University of Sussex, United Kingdom

[†] Current address: Department of Psychology, University of Portsmouth, United Kingdom

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

Running Head: Longitudinal study of voice pitch

Word count: 2533

Abstract

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 [1], 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.

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

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 [2] or pre-pubertal children's speech [3 for review]. In contrast, strong individual differences characterise 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 80-175 Hz and women's between 160-270 Hz [3]) and in their nonverbal vocalisations (e.g. tennis grunts: men 230-400 Hz, women 370-910 [4]). Among pre-pubertal children, in the absence of sex differences, *F*0 distributions largely overlap and are characterised by wide ranges in both preverbal cries (*F*0 ranges 350-550 Hz in the cries of 3-month-old babies [2]) and children's speech (215 to 320 Hz in 5 to 7-year-old's speech [5]).

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 [1]. 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.

Here, we test whether the *F*0 of preverbal 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 the *F*0 of their speech at 5 years of age. Because lower 2D:4D ratios have been linked to higher fetal 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.

Methods

Participants

We recorded the voices of 15 French children (6 females and 9 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 [2]).

Participants were recorded in their own home, 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.

Audio Recordings and Analyses

To obtain the $F0$ 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.5.2.1 [10]. Recordings were edited manually to remove all silences, nonverbal vocalisations (e.g., laughter, loud breathing or nonverbal interjections) acute background noises and the experimenter's questions, resulting in concatenated speech sequences ranging between 19.9 and 73.3 seconds (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 [2] 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 seconds (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.

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 3 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.

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.

Results

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).

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. Range *F*0 is reported for information.

	Sex (n)	Mean	(Range)	SE	*t*[1]	*df*	*p*
Baby cry F0	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			

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

Relationship between baby cry F0 and child speech F0

A stepwise multiple linear regression showed that neither the age, weight nor height at time of voice recording 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 =.637, p=0.011; Table 2, Figure 1).

Table 2. Stepwise multiple linear regression examining the effects of age and body size at five-year old, as well as *F*0 as a baby, on the mean *F*0 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

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

Relationships between 2D:4D ratio and F0

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 =.648, p=0.009; Figure 2). Inter-individual differences in children's right 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. In contrast, left hand 2D:4D ratios were not significant predictors of *F*0 at either age (both p>0.05). See Electronic Supplementary Materials (Table S1) for full descriptive correlation matrix.

Discussion

Pitch is a highly salient feature of the human voice that affects listeners' perceptions of femininity and masculinity in babies' cries [2], 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, preverbal 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 [1], 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 [2,5], and justifiying 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 two-years [6] and 6-14 years [17]) and often nonsignificant [6].

Given the small sample size in our investigation, constrained by the difficulty of obtaining longitudinal data on 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 if 2D:4D digit ratios are also predictors of voice pitch in adult speakers.

Figure Captions

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 5 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).

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).

Author contributions

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

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

Competing Interests

The authors report no competing interests.

Data accessibility

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

Acknowledgements

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

Funding statement

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

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)

References

1. 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.

2. Reby D, Levréro F, Gustafsson E, Mathevon N. 2016 Sex stereotypes influence adults' perception of babies' cries. *BMC Psychol.* **4**, 19.

3. Kreiman J, Sidtis D. 2011 *Foundations of voice studies: An interdisciplinary approach to voice production and perception*. Wiley-Blackwell.

4. Raine J, Pisanski K, Reby D. 2017 Tennis grunts communicate acoustic cues to sex and contest outcome. *Anim. Behav.* **130**, 47–55.

5. 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.

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.

7. Zheng Z, Cohn MJ. 2011 Developmental basis of sexually dimorphic digit ratios. *Proc. Natl. Acad. Sci.* **108**, 16289–16294.

8. Berenbaum SA, Bryk KK, Nowak N, Quigley CA, Moffat S. 2009 Fingers as a marker of prenatal androgen exposure. *Endocrinology* **150**, 5119–5124.

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.

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.

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.

14. Titze IR. 2011 Vocal fold mass is not a useful quantity for describing f0 in vocalization. *J. Speech Lang. Hear. Res.* **54**, 520–522.

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.

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.

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.

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. Lond. B Biol. Sci.*, rspb20112351.

Figure 1

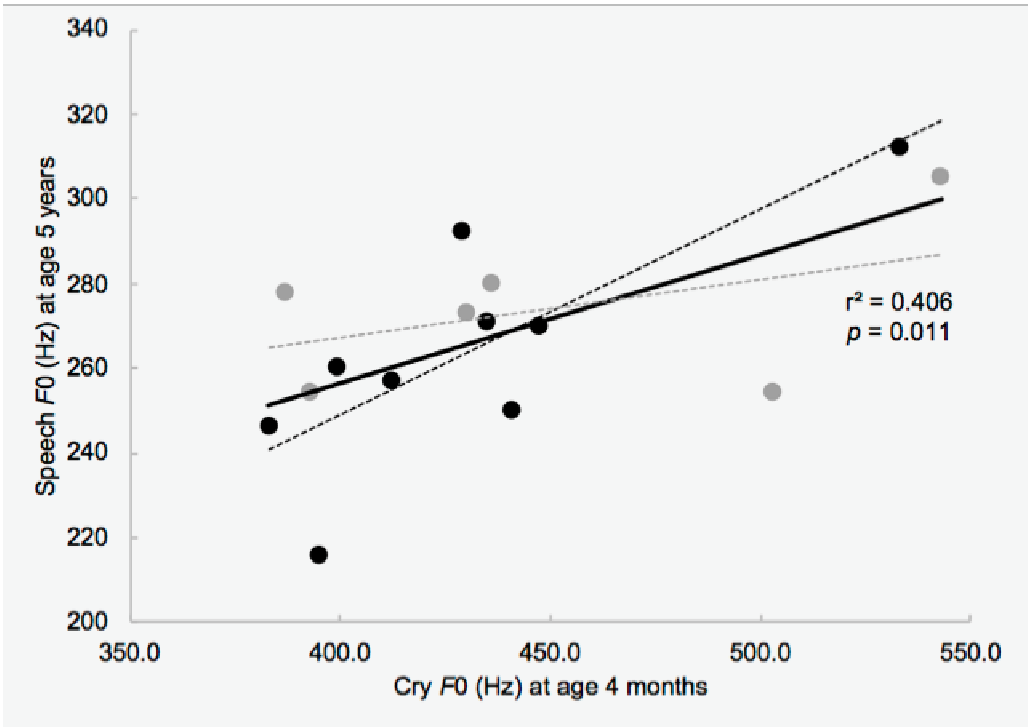
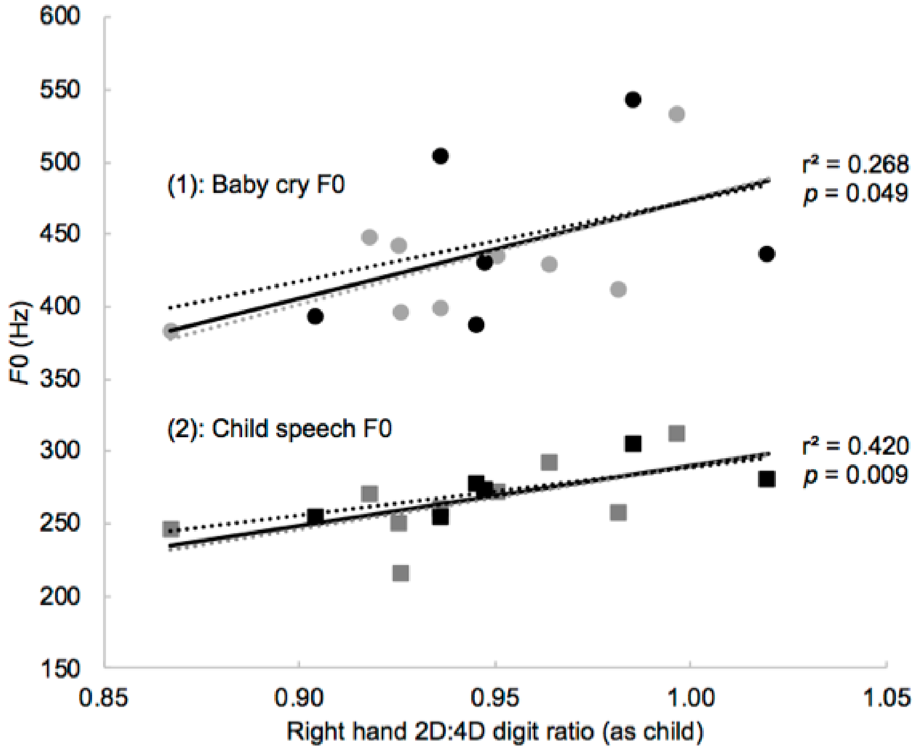


Figure 2


600
550
500
450
400
350
300
250
200
150
F0 (Hz)
(1): Baby cry F0
r² = 0.268
p = 0.049
(2): Child speech F0
r² = 0.420
p = 0.009
0.85
0.90
0.95
1.00
1.05
Right hand 2D:4D digit ratio (as child)