coordinated activity of these networks in the striatal function, and, therefore, they may participate in its dysfunction in brain disorders. Our results demonstrate the existence of functional astro-neuronal networks that comprise subpopulations of astrocytes, neurons, and synapses belonging to specific brain circuits, which may differentially control specific circuit activity through selective signaling between particular astrocytes and neurons.

REFERENCES AND NOTES

1. A. Araque *et al.*, *Neuron* **81**, 728–739 (2014).
2. C. Eroglu, B. A. Barres, *Nature* **468**, 223–231 (2010).
3. M. M. Halassa, P. G. Haydon, *Annu. Rev. Physiol.* **72**, 335–355 (2010).
4. G. Perea, M. Navarrete, A. Araque, *Trends Neurosci.* **32**, 421–431 (2009).
5. A. Volterra, J. Meldolesi, *Nat. Rev. Neurosci.* **6**, 626–640 (2005).
6. G. Perea, A. Araque, *Science* **317**, 1083–1086 (2007).
7. C. Agulhon, T. A. Fiacco, K. D. McCarthy, *Science* **327**, 1250–1254 (2010).
8. T. A. Fiacco *et al.*, *Neuron* **54**, 611–626 (2007).
9. B. C. Shonesy *et al.*, *Nat. Neurosci.* **16**, 456–463 (2013).
10. A. C. Kreitzer, R. C. Malenka, *Nature* **445**, 643–647 (2007).
11. W. Shen, M. Flajolet, P. Greengard, D. J. Surmeier, *Science* **321**, 848–851 (2008).
12. M. Uchigashima *et al.*, *J. Neurosci.* **27**, 3663–3676 (2007).
13. V. Chevaleyre, K. A. Takahashi, P. E. Castillo, *Annu. Rev. Neurosci.* **29**, 37–76 (2006).
14. A. C. Kreitzer, R. C. Malenka, *Neuron* **60**, 543–554 (2008).
15. R. L. Albin, A. B. Young, J. B. Penney, *Trends Neurosci.* **12**, 366–375 (1989).
16. L. M. Suárez *et al.*, *Biol. Psychiatry* **75**, 711–722 (2014).
17. S. Ares-Santos, N. Granado, I. Espadas, R. Martinez-Murillo, R. Moratalla, *Neuropsychopharmacology* **39**, 1066–1080 (2014).
18. M. Gómez-Gonzalo *et al.*, *Cereb. Cortex* (2014).
19. M. Navarrete, A. Araque, *Neuron* **68**, 113–126 (2010).
20. R. Min, T. Nevian, *Nat. Neurosci.* **15**, 746–753 (2012).
21. M. Navarrete *et al.*, *PLOS Biol.* **10**, e1001259 (2012).
22. M. A. Di Castro *et al.*, *Nat. Neurosci.* **14**, 1276–1284 (2011).
23. J. Petravicz, T. A. Fiacco, K. D. McCarthy, *J. Neurosci.* **28**, 4967–4973 (2008).
24. T. Fellin *et al.*, *Neuron* **43**, 729–743 (2004).
25. M. Navarrete, A. Araque, *Neuron* **57**, 883–893 (2008).
26. G. Perea, A. Araque, *J. Neurosci.* **25**, 2192–2203 (2005).
27. G. Cui *et al.*, *Nature* **494**, 238–242 (2013).
28. A. V. Kravitz *et al.*, *Nature* **466**, 622–626 (2010).
29. I. Ruiz-DeDiego, B. Mellstrom, M. Vallejo, J. R. Naranjo, R. Moratalla, *Biol. Psychiatry* **77**, 95–105 (2015).
30. V. M. André *et al.*, *J. Neurosci.* **31**, 1170–1182 (2011).

ACKNOWLEDGMENTS

We thank D. Redish, W. Buño, A. Díez, and R. Gómez for helpful comments; J. Lerma and A. Valero for valuable help; B. Pro for technical assistance; and J. Chen (University of California–San Diego) and A. Zimmer (University of Bonn) for providing $IP_3R2^{-/-}$ and $CB1R^{-/-}$ mice, respectively. A.A. was supported by Cajal Blue Brain, the Human Frontier Science Program (grant RGP0036/2014), the Wallin family, and V. and L. Han. Support for this work was also provided by the Ministerio de Economía y Competitividad of Spain (grants BFU2013-47265-R, CSD2010-00045, and RYC-2012-12014 to G.P. and grants SAF2013-48532-R and CB06/05/0055 to R.Mo.) and by the Juan de la Cierva Program (grant JCI-2010-07693 to R.Ma.). Data described in the paper are archived at the Instituto Cajal (Madrid, Spain). The authors declare no competing financial interests.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6249/730/suppl/DC1
Materials and Methods
Figs. S1 to S7
References (*31*–*39*)

28 January 2015; accepted 17 June 2015
10.1126/science.aaa7945

LANGUAGE DEVELOPMENT

The developmental dynamics of marmoset monkey vocal production

D. Y. Takahashi,[1,2]* A. R. Fenley,[1,2] Y. Teramoto,[1] D. Z. Narayanan,[1,2] J. I. Borjon,[1,2] P. Holmes,[1,3] A. A. Ghazanfar[1,2,4]*

Human vocal development occurs through two parallel interactive processes that transform infant cries into more mature vocalizations, such as cooing sounds and babbling. First, natural categories of sounds change as the vocal apparatus matures. Second, parental vocal feedback sensitizes infants to certain features of those sounds, and the sounds are modified accordingly. Paradoxically, our closest living ancestors, nonhuman primates, are thought to undergo few or no production-related acoustic changes during development, and any such changes are thought to be impervious to social feedback. Using early and dense sampling, quantitative tracking of acoustic changes, and biomechanical modeling, we showed that vocalizations in infant marmoset monkeys undergo dramatic changes that cannot be solely attributed to simple consequences of growth. Using parental interaction experiments, we found that contingent parental feedback influences the rate of vocal development. These findings overturn decades-old ideas about primate vocalizations and show that marmoset monkeys are a compelling model system for early vocal development in humans.

Human vocal development is the outcome of interactions among an infant's developing body and nervous system and his or her experience with caregivers (*1*, *2*). Infant cries decline over the first 3 months as they transition into preverbal vocalizations (*3*). The rates of these transitions are influenced by social feedback: Contingent responses of caregivers spur the development of more mature vocalizations (*4*). In contrast, nonhuman primate vocalizations are widely viewed as undergoing little or no production-related acoustic changes during development, and any such changes are attributed solely to passive consequences of growth (*5*).

We tracked the vocal development of marmoset monkeys (*Callithrix jacchus*; $n = 10$)—a voluble, cooperative breeding species (*6*)—from the first postnatal day (P1) until they produced adultlike calls at 2 months of age. Recordings were taken at least twice weekly in two contexts: undirected (social isolation) and directed (with auditory, but not visual, contact with their mother or father). Such early and dense sampling is necessary to accurately capture developmental changes in marmosets because this species develops rapidly (*7*). Each recording session began with ~5 min in the undirected context followed by ~15 min in the directed context, with mothers and fathers alternating between each session. In the undirected context, infants exhibited a dramatic change in vocal production (Fig. 1A and audio S1 to S8). At P1, vocalizations were more numerous and variable in their spectrotemporal structure than those recorded in later weeks. The number and variability of calls diminished over 2 months, approaching mature vocal output with exclusive production of whistle-like "phee" calls in this context (*8*).

To quantify this developmental change as a continuous process without the bias of ethological labels (*9*), for each of the 73,421 recorded utterances, we measured four acoustic parameters similar to those used for tracking birdsong development (*10*): duration, dominant frequency, amplitude modulation (AM) frequency, and Wiener entropy (a measure of spectral flatness) (Fig. 1B). Changes in all four parameters were statistically significant ($n = 301$ sessions, $P < 0.001$), showing that vocalizations underwent a transformation in the first 2 months, whereby utterances lengthened, dominant and AM frequencies decreased, and entropy decreased. This pattern of change is consistent with both human and songbird vocal development (*10*, *11*). These changes in infant vocalizations, although not subtle, may be due solely to physical maturation (*5*). To test this, we used body weight as a proxy for overall growth [weight correlates well with vocal apparatus size in monkeys (*12*)]. Weight changes visibly contrasted with the trajectories of the acoustic parameters (Fig. 1, B and C). To quantify this difference, we used weight to predict changes in the acoustic parameters. Predicted average parameter values, given the average weight for each postnatal day, are shown in Fig. 1B. If growth completely explained the acoustic change, the residues would be uncorrelated and identically distributed across postnatal days. Using the Akaike information criterion (AIC), the best polynomial-fit order was three

[1]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA. [2]Department of Psychology, Princeton University, Princeton, NJ 08544, USA. [3]Department of Mechanical and Aerospace Engineering and Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA. [4]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA.
***Corresponding author. E-mail: takahashiyd@gmail.com (D.Y.T.); asifg@princeton.edu (A.A.G.)**

for all residues related to the acoustic parameters (Fig. 1D). To account for possible nonlinear relationships between growth and acoustic parameters, we log-transformed the weight and acoustic parameters. The log-transformed weight did not predict the log-transformed acoustic parameters (fig. S1, A to C). Thus, simple patterns of growth (linear or nonlinear) do not accurately predict acoustic changes in infant marmoset vocalizations.

A subset of the early vocalizations of humans and songbirds are incorporated into the adult repertoire, whereas others are transient, serving as scaffolding for later vocalizations (*3*, *10*). To test whether infant marmosets follow a similar trajectory, we first measured the extent to which their calls were distinct. Two parameters, duration and entropy, identified disjoint clusters in syllable sequences (Fig. 2A). With development, the clusters became more distinct and less numerous. Using all four parameters, we computed optimal cluster numbers for each marmoset in each session (Fig. 2B). On average, the number of clusters decreased from around four to one or two ($P < 0.001$). The clusters represent distinct ethologically based syllable types (*13*, *14*) (Fig. 2C). Phee syllables increased to over 95% of all vocalizations by 2 months ($P < 0.001$); all other calls decreased ($P = 0.005$ for trills; $P < 0.001$ for all other syllables) (Fig. 2D). The changes in syllable proportions potentially represent two independent processes: change in usage (*15*) and transformation of immature calls into mature versions (Fig. 2E). Twitters and trills are produced frequently by marmosets of all ages (*13*, *14*), but in adults, they are typically produced when in visual contact with conspecifics and not in the undirected context. Thus, twitters and trills undergo a change in usage in the first 2 months. In contrast, cries, phee-cries, and subharmonic phees are only produced by infants; mature phees are produced almost exclusively during vocal exchanges that occur when out of visual contact with conspecifics (*8*).

Because these infant-only calls share some features with the mature phee call (e.g., a common duration), we hypothesized that they represent immature phees, consistent with vocal transformations observed in preverbal human infants (*11*) and songbirds (*10*) but contrasting with previous reports on developing primates (*5*). It is possible that these transitional forms are related to growth but sound distinct because of nonlinearities in the vocal production

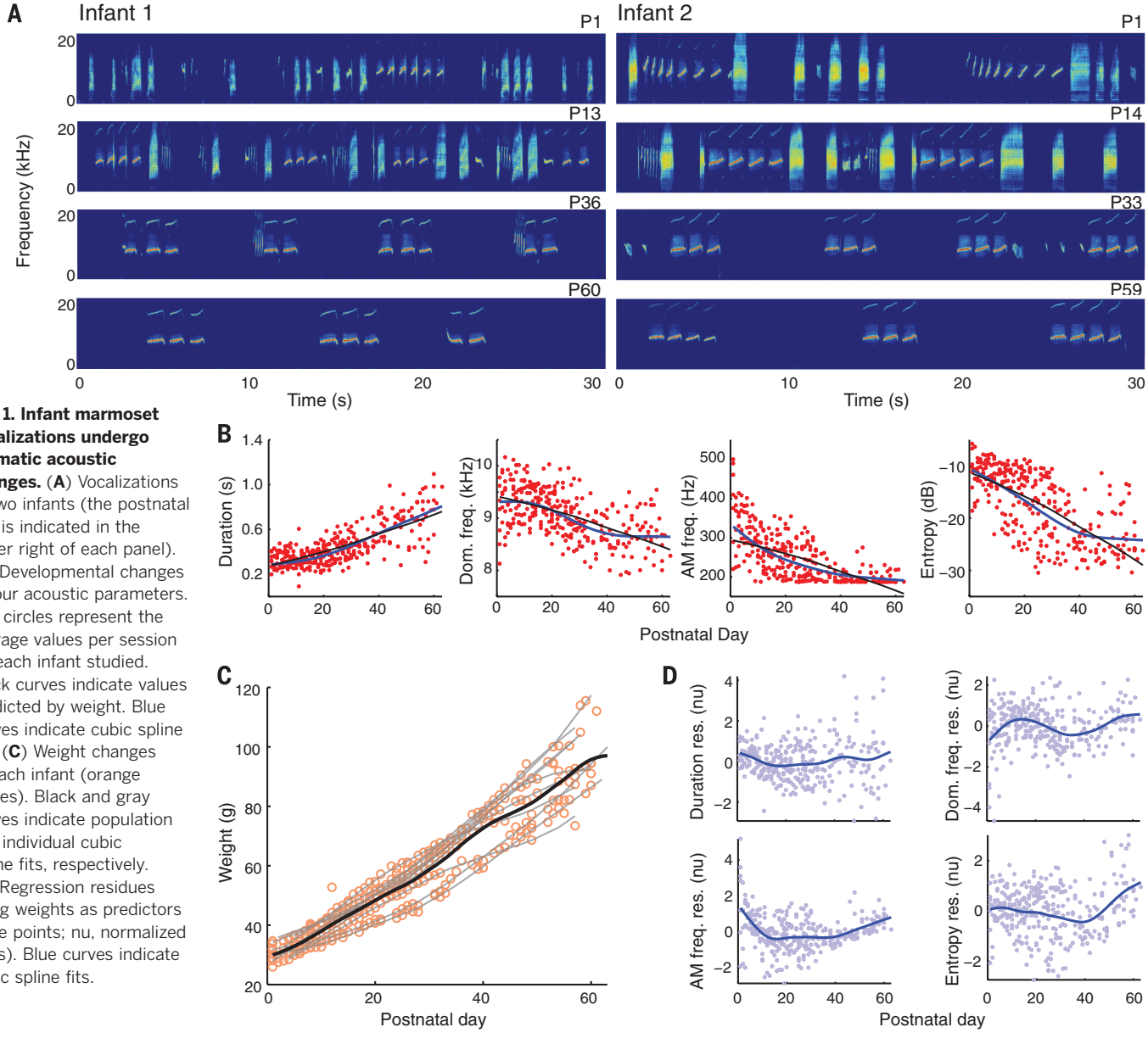

Fig. 1. Infant marmoset vocalizations undergo dramatic acoustic changes. (**A**) Vocalizations of two infants (the postnatal day is indicated in the upper right of each panel). (**B**) Developmental changes in four acoustic parameters. Red circles represent the average values per session for each infant studied. Black curves indicate values predicted by weight. Blue curves indicate cubic spline fits. (**C**) Weight changes of each infant (orange circles). Black and gray curves indicate population and individual cubic spline fits, respectively. (**D**) Regression residues using weights as predictors (blue points; nu, normalized units). Blue curves indicate cubic spline fits.

system (*16*). This would suggest that a single biomechanical mechanism generates cries, phees, and the transitional forms and that the transitional calls result from smooth changes through a parameter space. To test this idea, we developed a model based on one that successfully reproduces syllable types in zebra finch song but that is nonspecific with regard to songbird versus mammalian vocal anatomy (Fig. 3A) (*17*). Our simulations verified that the model can reproduce the marmoset call types described above (Fig. 3, B to E). The simulations also revealed the underlying biomechanics corresponding to different calls at different levels of pressure (respiratory power) and laryngeal muscle tension (Fig. 3F). Broadband cries were produced at low pressure and muscle tension, where small variations cause large changes in spectral content because of nonlinear vocal fold dynamics. Phees occurred at higher pressures and tensions, and subharmonic phees occurred in an intermediate region, supporting their classification as transitional calls. Rapid switching between high and low pressure and tension states produced the phee-cries. Throughout, linear changes in pressure and tension produced nonlinear acoustic effects.

To test the model's overall validity and the prediction that respiration during cries is less stable than during phees, we measured respiratory activity via electromyography (EMG) in five P1 infants. We investigated whether different respiratory patterns underlie cries and phees with similar intersyllable intervals (Fig. 3, G and I). The EMG signals were more uniform across phees than across cries (Fig. 3, H and J), as quantified by the cost of dynamic time warping (DTW) (*18*). For each infant, the mean DTW costs for phees were smaller than they were for cries ($P < 0.001$) (Fig. 3K). Therefore, phee syllables at least partly result from more stable respiration; immature respiratory control leads mainly to cries early in life, consistent with the model prediction. Overall, these data support our hypothesis that cries are immature phees.

Thus, although vocal acoustic changes were dramatic, physiological growth could explain the transition from cries to phees, as improved respiratory and/or laryngeal control modulates spectral parameters (Fig. 1B), reducing the entropy. However, if the cries-to-phees transition was solely driven by physical maturation, it would be impervious to social feedback. Yet, consistent with a role for vocal feedback in guiding development, marmoset monkeys exhibit a developmental pattern of *FoxP2* expression in their thalamocortical-basal ganglia circuit (*19*) that is analogous to that of songbirds and humans (*20*). This suggests that marmoset infants may use this circuit to guide their phee-call development through reward-based parental feedback, as birds and humans do (*21*). To assess the effect of parent-infant vocal interactions in marmosets, we quantified their vocal exchanges in the directed context, where infants and their mother or father were in auditory, but not visual, contact.

Infant and parent vocalizations were parsed into whole multisyllabic calls according to the

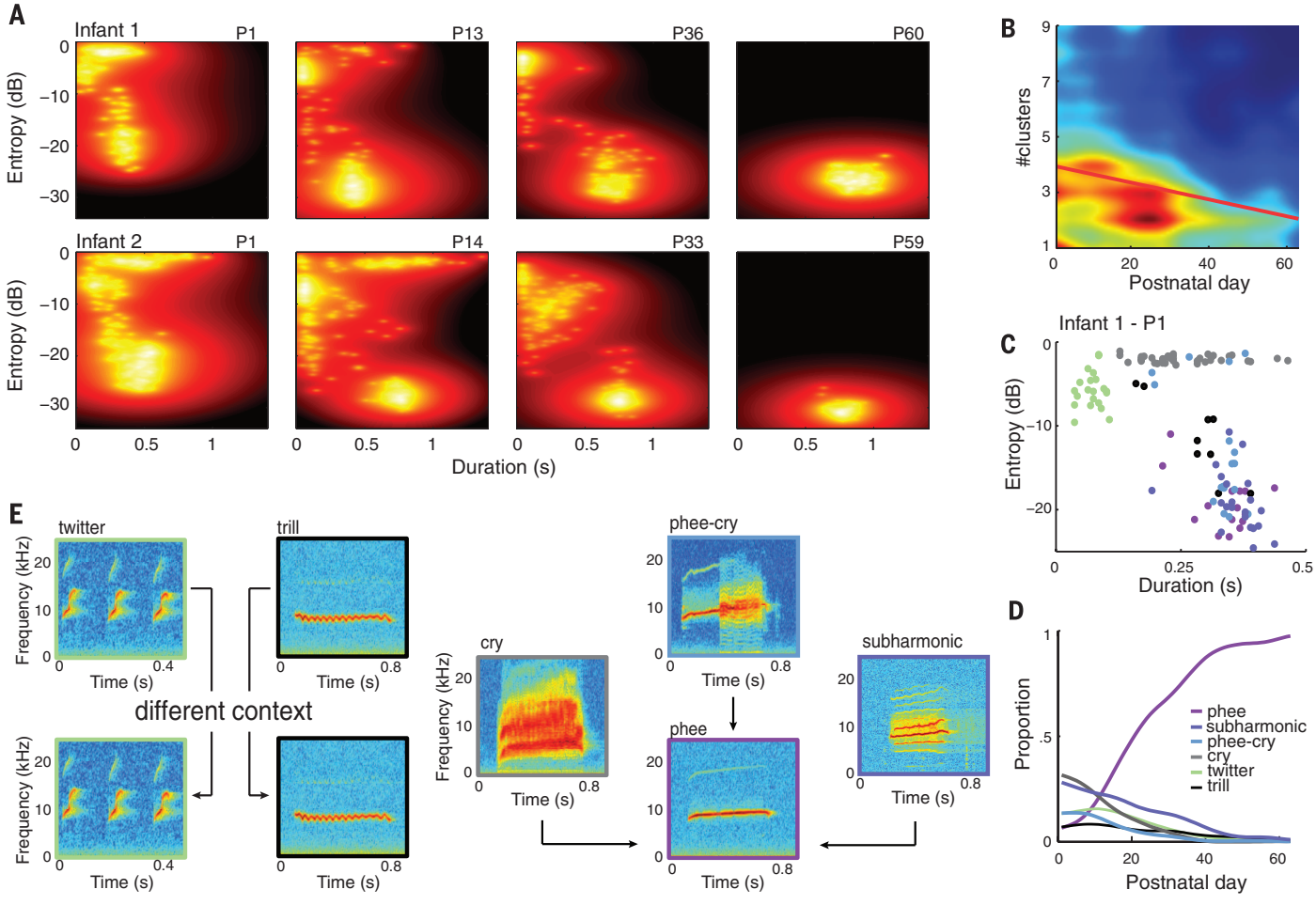

Fig. 2. A subset of infant marmoset calls transform into adultlike phee calls. (**A**) Probability maps of duration and entropy for the infants in Fig. 1A. Lighter colors indicate higher probabilities. (**B**) Distribution of cluster numbers by developmental age. Warmer colors indicate more frequent occurrences of cluster numbers. The red line is the regression fit. (**C**) Correspondence between clusters and different syllable types for a marmoset at P1. (**D**) Developmental changes in the proportion of call types. Colors in (C) indicate the same call types as in (D). (**E**) Twitters and trills change in usage, whereas cries, phee-cries, and subharmonic phees are hypothesized to transition to phee calls.

bimodal distribution of their intersyllable intervals (*8*). We recorded 8800 infant phees, 11,798 infant cries, and 6567 adult phees, of which 2512 were contingent responses to infant phees [those falling within a turn-taking interval as seen in adults (*8*)]. Parents produced mostly phee calls (>98%). Typical examples of infant phee and cry production during interactions over the first 2 months and the phee/cry ratio across days are shown in Fig. 4, A and B. As in the undirected context (Fig. 2D), cries gave way to phees, but the transition occurred rapidly. For each infant, we used the point where the phee/cry ratio first crossed zero to mark the transition day (Fig. 4C). Transitions were typically sharp, but their timing varied substantially across infants (~10 to 40 postnatal days). If physiological growth completely explained the cries-to-phees transition, the weight-change rate and the timing of the transition (zero-crossing) day would be correlated. However, we found no significant correlation (n = 10 infants, t test, P = 0.684) (Fig. 4D); growth alone cannot explain the timing of the cries-to-phees transition.

We then investigated whether parental responses to infant vocalizations affect the timing of the cries-to-phees transition. This would explain, at least partially, its variability across infants. Infants could be influenced by contingent responses only or by the total number of adult vocalizations that they hear. The fraction of infant phees that elicited contingent parental phee responses before the zero-crossing day correlated significantly with the timing of the zero-crossing day (n = 10 infants, t test, P = 0.005) (Fig. 4E). Proportions of noncontingent parental calls (91.5% of all calls on average) were not significantly correlated with this timing (n = 10 infants, t test, P = 0.558) (Fig. 4F). Therefore, contingent vocal responses from parents influence the timing of the cries-to-phees transition by reinforcing the production of phee calls.

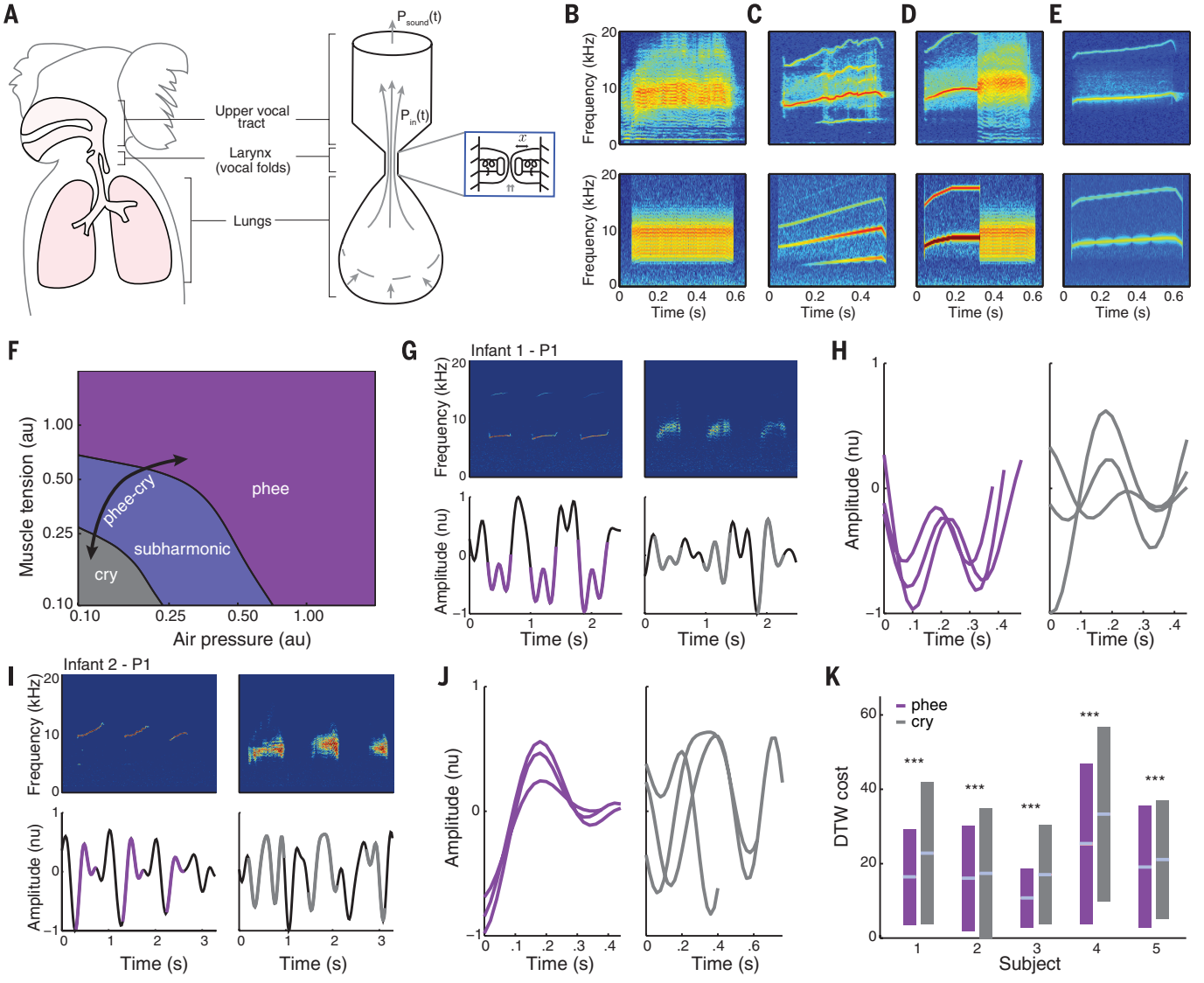

Fig. 3. Biomechanical model of vocal folds can reproduce infant marmoset calls. (**A**) Main elements of the vocal tract and respective model representations, where $P_{in}(t)$ is pressure at the inlet to the vocal tract, $P_{sound}(t)$ is pressure upon exit from the mouth, and × is the displacement of the vocal fold. Gray arrows show airflow. (**B** to **E**) Top: Representative recordings of a cry (B), subharmonic phee (C), phee-cry (D), and phee (E); bottom: corresponding model simulations. (**F**) Changing air pressure and muscle tension produces different calls (au, arbitrary units). (**G** and **I**) Top: Three-syllable phees (left) and cries (right) for two marmosets; bottom: their corresponding EMG activities. (**H** and **J**) Respiratory EMG activity during call production. Curves are the purple and gray segments from (G) and (I), aligned at the syllable onset of phees and cries, respectively. (**K**) Average DTW cost on P1 for five marmosets. Central light-colored bars show means, rectangles show standard deviations, and asterisks indicate significant differences (P < 0.001).

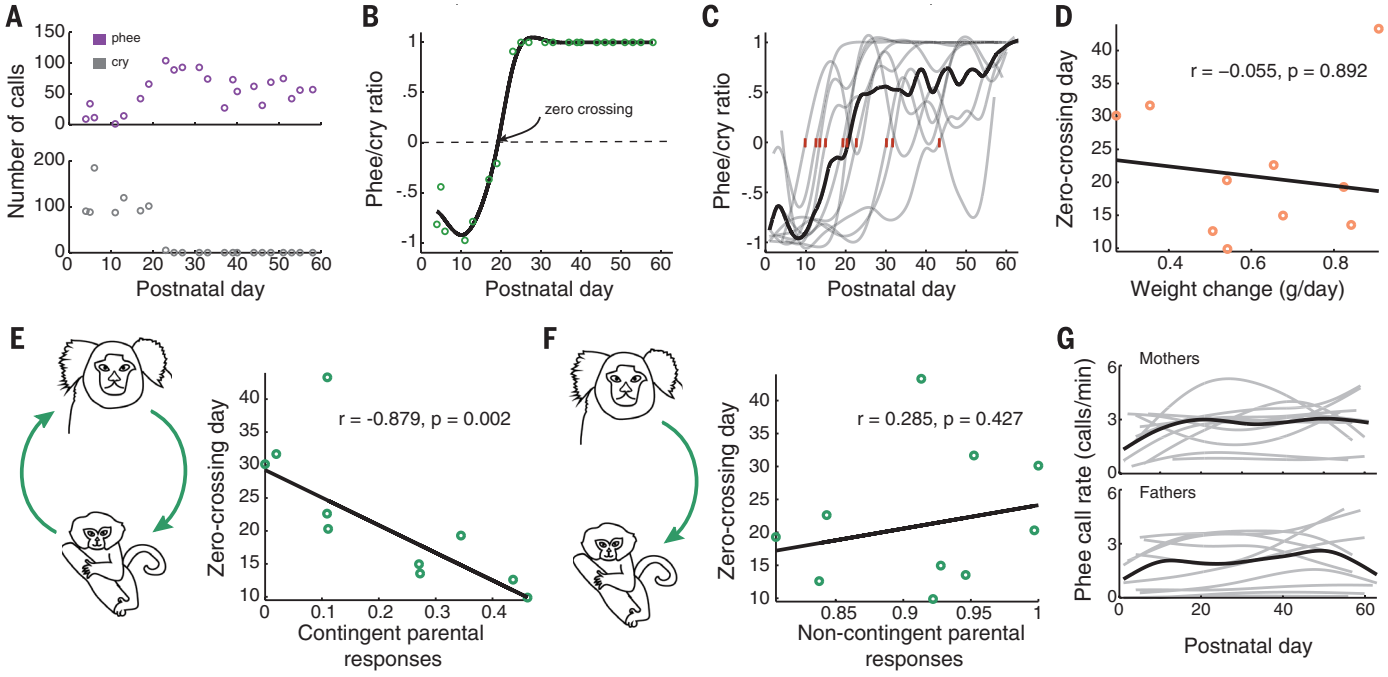

Fig. 4. Transition from cry to phee is influenced by contingent parental calls. (**A**) Numbers of cries and phees over 2 months for a single infant. (**B**) Phee/cry ratio for the infant in (A) across days. (**C**) Phee/cry ratios (gray curves) and zero-crossing days (red ticks) for each infant and for the population (black curve). Black and gray curves in (B) and (C) are cubic spline fits. (**D**) Correlation between the weight-change rate and the zero-crossing day among infants. (**E** and **F**) Correlations between the zero-crossing day and the proportion of contingent and noncontingent parental responses, respectively. (**G**) Rates of individual parental phee-call production during infant development (gray) and the population average (black).

We address two possible caveats to this conclusion. First, it is possible that, through shared genetics, fast-transitioning infants are born to more vocally interactive parents. To test this, we correlated the frequency of contingent parental calls and the zero-crossing day for six full siblings born from the same parents. If shared genetics were driving the result, then there would be no correlation between contingent parental responses and the zero-crossing day. We found, however, that there remained a statistically significant correlation (n = 6 infants, P = 0.046) (fig. S2). Moreover, we found no difference between the slopes of the regressions for the full-siblings and all-infants data (test for equality, P = 0.953).

Second, it is possible that changing patterns of infant calling are due to changes in parental call output. The phee-call production rates of each infant's parents during development are shown in Fig. 4G; neither parent changed their production rates (mother, P = 0.132; father, P = 0.235). Based on these analyses, we conclude that the cries-to-phees transition is influenced by contingent responses from parents, not by shared genetics or changes in parental vocal output.

Our findings demonstrate that infant marmoset calls undergo dramatic changes during the first 2 months of life, transforming from cries into mature, adultlike phee calls. The timing of this transition is partly attributable to maturation but is also influenced by contingent parental vocal feedback. This is consistent with preverbal vocal development in humans, whereby (i) natural categories of sounds change as respiratory, laryngeal, and facial components mature, and (ii) in parallel, vocal feedback sensitizes infants to certain features of those sounds, and the sounds are modified accordingly. Our findings contrast with previous reports that nonhuman primate vocalizations undergo little or no postnatal change and are impervious to social feedback (*5*). The complex and socially dependent vocal development we observed in marmoset monkeys may be a necessary condition of the vocal learning observed in humans.

REFERENCES AND NOTES

1. E. Thelen, in *Biological and Behavioral Determinants of Language Development*, N. A. Krasnegor, D. M. Rumbaugh, R. L. Schiefelbusch, M. Studdert-Kennedy, Eds. (Lawrence Erlbaum Associates, Hillsdale, NJ, 1991), pp. 339–362.
2. L. Byrge, O. Sporns, L. B. Smith, *Trends Cogn. Sci.* **18**, 395–403 (2014).
3. D. K. Oller, *The Emergence of the Speech Capacity* (Lawrence Erlbaum Associates, Mahwah, NJ, 2000).
4. M. H. Goldstein, A. P. King, M. J. West, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 8030–8035 (2003).
5. S. E. R. Egnor, M. D. Hauser, *Trends Neurosci.* **27**, 649–654 (2004).
6. J. I. Borjon, A. A. Ghazanfar, *Brain Behav. Evol.* **84**, 93–102 (2014).
7. A. de Castro Leão, A. Duarte Dória Neto, M. B. C. de Sousa, *Comput. Biol. Med.* **39**, 853–859 (2009).
8. D. Y. Takahashi, D. Z. Narayanan, A. A. Ghazanfar, *Curr. Biol.* **23**, 2162–2168 (2013).
9. D. Lipkind, O. Tchernichovski, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15572–15579 (2011).
10. O. Tchernichovski, P. P. Mitra, T. Lints, F. Nottebohm, *Science* **291**, 2564–2569 (2001).
11. E. Scheiner, K. Hammerschmidt, U. Jurgens, P. Zwirner, *J. Voice* **16**, 509–529 (2002).
12. W. T. Fitch, *J. Acoust. Soc. Am.* **102**, 1213–1222 (1997).
13. B. M. Bezerra, A. Souto, *Int. J. Primatol.* **29**, 671–701 (2008).
14. A. L. Pistorio, B. Vintch, X. Wang, *J. Acoust. Soc. Am.* **120**, 1655–1670 (2006).
15. A. M. Elowson, C. T. Snowdon, C. Lazaro-Perea, *Behaviour* **135**, 643–664 (1998).
16. W. T. Fitch, J. Neubauer, H. Herzel, *Anim. Behav.* **63**, 407–418 (2002).
17. A. Amador, Y. S. Perl, G. B. Mindlin, D. Margoliash, *Nature* **495**, 59–64 (2013).
18. J. A. Kogan, D. Margoliash, *J. Acoust. Soc. Am.* **103**, 2185–2196 (1998).
19. M. Kato *et al.*, *Brain Lang.* **133**, 26–38 (2014).
20. I. Teramitsu, L. C. Kudo, S. E. London, D. H. Geschwind, S. A. White, *J. Neurosci.* **24**, 3152–3163 (2004).
21. S. Syal, B. L. Finlay, *Dev. Sci.* **14**, 417–430 (2011).

ACKNOWLEDGMENTS

We thank L. Kelly for providing comments on an earlier draft. This work was supported by a Pew Latin American Fellowship (D.Y.T.), a Brazilian Science Without Borders Fellowship (D.Y.T.), an NSF Graduate Fellowship (J.I.B.), and the James S. McDonnell Scholar Award (A.A.G). Data for each figure are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6249/734/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 and S2
References (*22*–*29*)
Audio S1 to S8
Supplementary Data

13 March 2015; accepted 12 June 2015
10.1126/science.aab1058

www.sciencemag.org/content/349/6249/734/suppl/DC1

Supplementary Materials for

The developmental dynamics of marmoset monkey vocal production

D. Y. Takahashi, * A. R. Fenley, Y. Teramoto, D. Z. Narayanan, J. I. Borjon, P. Holmes, A. A. Ghazanfar*

*Corresponding author. Email: takahashiyd@gmail.com (D.Y.T.); asifg@princeton.edu (A.A.G.)

Published 14 August 2015, *Science* **349**, 734 (2015)
DOI: 10.1126/science.aab1058

This PDF file includes:

Other Supporting Online Material for this manuscript includes the following:
(available at www. sciencemag.org/content/349/6249/734/suppl/DC1)

Materials and Methods

Subjects

All experiments were performed in compliance with the guidelines of the Princeton University Institutional Animal Care and Use Committee. The subjects used in the study were 15 infants and 6 adults (3 male-female pairs, > 2 years old), captive common marmosets (*Callithrix jacchus*) housed at Princeton University. Ten infants (all members of twin sets with three twin sets from the same parents) participated in the undirected and directed calls experiment during the first two months of postnatal period. Another five infants (one set of twins, one set of triplets) were used for respiratory electromyography (EMG) experiments on the first postnatal day. The colony room is maintained at a temperature of approximately 27°C and 50-60% relative humidity, with 12L:12D light cycle. Marmosets live in family groups; all were born in captivity. They had *ad libitum* access to water and were fed daily with standard commercial chow supplemented with fruits and vegetables. Additional treats (peanuts, cereal, dried fruits and marshmallows) were used prior to each session to transfer the animals from their home-cage into a transfer cage.

Experimental setup

Beginning on their first postnatal day, we recorded the vocalizations of marmoset monkey infants in two different contexts: undirected (i.e., social isolation) and directed (with auditory, but not visual, contact with their mother or father). Early in life, infants are always carried by parent. Thus, the parent carrying the infant(s) was first lured from the home cage into a transfer cage using treats. The infant marmoset was then gently separated from the adult and taken to the experiment room where it was placed in a second transfer cage on a flat piece of foam. Avoiding separation longer than 30 minutes in one day and alternating with shorter sessions for undirected experiments (~ 5 min) minimized the stress caused by separation to the infant. The vocalizations we observed were identical in type to those produced when the infant is naturally separated from parents (e.g., when parents push them off or when they transfer them to the other parent for carrying or feeding). The cage rested on a table (.66m in height) in one of two opposing corners of the room. The testing corner was counterbalanced across sessions. A speaker was placed at a third corner equidistant from both testing corners and pink noise (amplitude decaying inversely proportional to frequency) was broadcast at ~45 dB (at 0.88m from speaker) in order to mask occasional noises produced external to the testing room. An opaque curtain made of black clothes divided the room to visually occlude the subject from the other corner. A digital recorder (ZOOM H4n Handy Recorder) was placed directly in front of the transfer cage at a distance of .76m. Audio signals were acquired at a sampling frequency of 96kHz. Every session typically consisted of two consecutive undirected experiments (one twin followed by the other) and one directed experiment (just one of the twins on a given day). Each session started with the undirected experiments lasting ~5 minutes each one. The number of undirected experiments with at least one call production was 40, 38, 38, 38, 37, 39, 19, 15, 16, 21 (10 infants, 301 sessions, 73,421 utterances). The order of the infants was counterbalanced. As soon as the undirected experiment was finished, one of the parents

was brought to the experiment room and put into the opposing corner of the room. A second digital recorder (ZOOM H4n Handy Recorder) was placed directly in front of the parent at a distance of .76m from the transfer cage. During this setup procedure and throughout the directed experiment, the opaque curtain prevented the infant and the parent from having visual contact. The directed experiment lasted for ~15 min. The order of which parent participated in the interaction was counterbalanced. If the parent took more than 15 minutes to be lured for the directed calls experiment, the experiment was aborted to avoid any excessive separation stress on infants and parents. The number of directed experiments for each infant was 17, 13, 13, 18, 24, 24, 22, 21, 21, 22 (195 sessions).

Detection of calls

To determine the onset and offset of a syllable, a custom made MATLAB® routine automatically detected the onset and offset of any signal that differed from the background noise at specific frequency range. To detect the differences, we first bandpass filtered the entire recording signal between 6 and 10kHz. This corresponds to the dominant frequency of marmoset calls, *i.e.*, the frequency with highest power, which is not necessarily the fundamental frequency (F0), *i.e.*, the lowest frequency of the periodic components of the sound. Second, we resampled the signal to 1kHz sampling rate, applied the Hilbert transform and calculated the absolute value to obtain the amplitude envelope of the signal. The amplitude envelope was further low pass filtered to 50Hz. A segment of the recording without any call (silent) was chosen as a comparison baseline. The 99th percentile of the amplitude value in the silent period was used as the detection threshold. Sounds with amplitude envelope higher than the threshold were considered a possible vocalization. Finally, to ensure that sounds other than call syllables were not included, a researcher verified whether each detected sound was a vocalization or not based on the spectrogram.

Quantification of acoustic parameters

After detecting the onset and offset of calls, a custom made MATLAB® routine calculated the duration, dominant frequency, amplitude modulation frequency, and Wiener entropy of each syllable. The duration of syllable is the difference between the offset and onset of the vocalization detected by our custom made program discussed above. The dominant frequency of a syllable was calculated as the average frequency at which the spectrogram had maximum power. The spectrogram was calculated using a FFT window of 1024 points, Hanning window, with 50% overlap. The amplitude modulation frequency was calculated in the following way. First, the signal was bandpass filtered between 6 to 10kHz and then a Hilbert transform was applied. The absolute value of the resulting signal gives us the amplitude envelope of the modulated signal. Finally, the amplitude modulation frequency was calculated as the dominant frequency of the amplitude envelope. The Wiener entropy is the logarithm of the ratio between the geometric and arithmetic means of the values of power spectrum for different frequencies (*10*). The Wiener entropy represents how broadband the power spectrum of a signal is. The closer the signal is to white noise, the higher will be the value of Wiener entropy. A cubic spline curve was fitted to the population data using MATLAB® csaps function. To verify if the parameters changed during development, we fitted a robust linear regression

(robustfit) using MATLAB® robustfit. We used the two-sided t-test for the nullity of the slope to verify the statistical significance of the slope of the linear regression (n = 301 sessions). A cubic spline curve was fitted to individual and population weight to show the almost linear growth curve.

To test if body weight, which is a highly correlated to vocal tract length (*12*), can predict the observed acoustic changes, we fitted a robust linear regression to each of the acoustic parameter using the weight as predictor. The robust linear regression and the respective residuals were calculated for each infant separately (n = 45, 42, 43, 36, 37, 37, 14, 12, 13, 16 days of weight measurements for each infant). A cubic spline curve was fitted to the all residuals for all infants. To test for presence of nonlinearity in the residuals, we used Akaike Information Criterion (AIC) to select the order of the best polynomial fit on the residuals. We used the polydeg routine to calculate AIC (http://www.biomecardio.com/matlab/polydeg.html). To obtain the predicted average population values for each acoustic parameter, we calculated the robust linear regression between the parameter values and weights for the population data. Then, we plotted the parameter value predicted for the average population weight for each postnatal day. To take into account possible polynomial nonlinearity in the relationship between the weight and acoustic parameters we applied the log transform to the weight and the acoustic parameters. The Wiener entropy is negative valued, therefore, we applied the log transform to the absolute value of the Wiener entropy. If the variables are related by some polynomial equation (e.g., $y = a*x^p$), the log transform will linearize the relationship (e.g., $\log(y) = p \log(x) + \log(a)$) and then standard linear statistical inference can be applied. Once the data was log transformed, we repeated the same procedure applied to non-transformed data to test if body weight can predict the observed acoustic changes.

Clustering analysis of the acoustic parameters

To calculate the number of clusters for each session and subject, we used the spectral clustering algorithm (*22*) on the four acoustic parameters (duration, dominant frequency, amplitude modulation frequency, Wiener entropy). We used the implementation by Ingo Bürk (MATLAB® file exchange #34412) for the spectral clustering analysis. Sessions with less than 20 calls were excluded because clustering algorithm was not reliable (259 session were included in the analysis). To determine the optimal number of clusters, we used slope statistics (*23*). We calculated the probability distribution of the optimal number of clusters chosen by slope statistics for all sessions and infants to show how the number of clusters changes during development. We fitted a robust linear regression to the number of clusters versus postnatal day. The statistical significance of the slope of the linear regression was measured using the two-sided t-test for the nullity of the regression slope (n = 259 sessions).

Classification of type of call syllables

Each automatically detected call was manually classified as phee, phee-cry, subharmonic-phee, cry, twitter, and trill, based on the spectro-temporal profile measured by the spectrogram. To ensure validity of our classification procedure, 10 sessions chosen at random were classified by two different individual and compared. The classification matched in more than 99.9% of the call syllables. The six call types show very distinct spectro-temporal profiles and can be easily classified by eye (*13, 14*). Briefly, phee is a

tonelike long call with F0 at around 7-10kHz; twitter is a short upward FM sweep; trill is defined by sinusoidal-like FM throughout the entire call; cry is a broad-band call, with F0 around 600Hz; phee-cry is a combination of phee and cry in any order, with each component lasting at least 50ms. A subharmonic-phee is similar to phee, but with a strong harmonic component around 3.5-5kHz. We classified a call as subharmonic-phee if the harmonic component around 3.5-5kHz is visible for at least 50 ms. For each session and subject, the proportion of each type of call (Fig. 2D) was calculated as the sum of all durations of each type of syllable divided by the sum of all durations of phee, phee-cry, subharmonic-phee, cry, twitter, and trill. We calculated the cubic spline curve for the development of the proportion of each type of syllable using MATLAB® csaps. To statistically test the significance of the developmental change (increase or decrease), we fitted a robust linear regression and applied the two-sided t-test to verify the nullity of regression slope.

Biophysical model of vocal production

To model vocal production, we adapted and extended a biomechanical model that was introduced to study song generation in zebra finch (*17*). This was derived from previous models developed to understand the biomechanics of human speech (*24*), using nonlinear coordinate changes to produce a simple "normal form" (*25*). With appropriate parameter choices, detailed below, the resulting model is adequate to describe primate vocal production. The vocal production apparatus is simplified into three parts: the respiratory system, the vibratory system (syrinx in birds, larynx in primates), and the filtering/resonance system (the supra-glottal vocal tract). The respiratory and vibratory systems are reduced to the following differential equations that describe the displacement x and velocity y of the vocal folds

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= -\alpha(t)\gamma^2 - \beta(t)\gamma^2 x - \gamma^2 x^3 - \gamma x^2 y + \gamma^2 x^2 - \gamma x y, \end{aligned} \qquad (1)$$

where $\gamma > 0$ is a time constant and two additional dimensionless parameters $\alpha(t)$ and $\beta(t)$, that may vary with time (t) or remain constant, respectively represent sub-glottal air pressure and laryngeal muscle tension. This single mass model, like that of (*24*), assumes that the vocal folds move symmetrically and support a traveling wave of fixed shape. We note that particular relationships among passive nonlinear stiffness and dissipation parameters must be assumed to obtain the form (1) with only 3 parameters.

Vocal fold vibrations are then translated into sound pressure changes in the supra-glottal vocal tract using the equation

$$P_{in}(t) = c_1 x(t) + c_2 \dot{x}(t) - c_3 \ddot{x}(t) - r P_{in}(t - T), \qquad (2)$$

where c_1, c_2, c_3 are positive constants and T is the time taken for sound to travel through the supra-glottal vocal tract to the mouth and, after reflection, back to the vocal folds. The constants c_j are coefficients of the leading terms of a Taylor series expansion for the incoming pressure as a function of flow velocity, as determined from vocal fold displacement, velocity and acceleration (*25*). The reflection coefficient $r < 1$ describes the amplitude change in the reflected pressure wave $rP_{in}(t - T)$ that returns to interact with the incoming signal $P_{in}(t)$. The resulting sound pressure emitted at the mouth is therefore

$$P_{sound}(t) = (1 - r)P_{in}\left(t - \frac{T}{2}\right). \quad (3)$$

Finally, we high pass filter the sound at 5kHz to simulate the filtering property of the supra-glottal vocal tract. Unlike the zebra finch work (*17*), we do not model the vocal cavity. The model parameters and their values used in this study are summarized in Table 1.

Table 1. Parameter values used for simulations. Summary of parameter values used to fit marmoset calls. The notation [0,2] means that values are chosen in the range 0 to 2.

Parameter	Interpretation	Values
dt	Simulation time step (μs)	5
α	Nondimensional air pressure (au)	[0, 2]
β	Nondimensional muscle tension (au)	[0, 2]
γ	Time constant (1/ms)	45
c_1, c_2, c_3	Pressure coefficients	(1, 0.1, 0.001)
r	Reflection coefficient	0.8
T/2	Time steps to travel down the vocal tract	9

To generate the simulated calls in Fig. 3, we varied the parameters $\alpha(t)$ and $\beta(t)$ within the range [0,2] indicated in Table 1 and matched the frequency spectra and temporal profiles of the simulated sound to the corresponding vocalizations. Numerical simulations of Equations (1-3) were carried out using Euler's method in custom written MATLAB® codes. To improve the fit between the model and recordings, pink noise was added to the simulation to better match its presence in the background of the exemplar vocalizations in Figs. 3B-E, using the MATLAB® pinknoise function (file exchange #42919 by Hristo Zhivomirov) (*26*). The parameter β was held fixed for the cry (Fig. 3B) while $\alpha(t)$ was ramped up and down in a piecewise linear manner; for the other calls, both $\alpha(t)$ and $\beta(t)$ were ramped up and down to produce the varying fundamental and harmonic frequencies of calls in Figs. 3C-E. High pass filtering of $P_{sound}(t)$ was done with MATLAB® eegfilt.

To separate the parameter space into cry, subharmonic-phee, and phee regions (Fig. 3F), we used the relationship between F0 and the natural frequency of the resonator. In our model, the natural frequency was 8kHz. If F0 was the same as the natural frequency, the simulated call was classified as phee. If F0 was half of the natural frequency, the simulated call was classified as a subharmonic-phee. If F0 was less or equal than one third of natural frequency, the call was classified as a cry.

Respiratory electromyography (EMG) signal

Infant marmosets (n=5 infants) were gently separated from the adult and taken to the experiment room where they were placed in a testing box that rested on a table (.66m in height) in one corner. Each recording session lasted for ~15 minutes. The testing box, made of plexiglas and wire, was in a triangular prism shape (.30m x .30m x .35m). To record the EMG signal, we used two pairs of Ag-AgCl surface electrodes (Glass technology). We put one pair of electrodes on the chest, close to the heart, and we put a second pair of electrodes on the back, close to the diaphragm. To improve the signal-to-noise ratio, we applied an ECl electrode gel on the surface of the electrode. Because the signal closer to the diaphragm showed the clearest respiratory signal, we used only EMG signals from the back for our analyses. Each pair of electrodes was differentially amplified (250x) and the resulting signal was sent to Plexon® omniplex, which digitized the signal at 40kHz and sent it to a PC. To obtain the respiratory EMG signal, we downsampled the recorded signal to 50Hz, bandpass filtered between 1 to 4Hz, and calculated the first derivative (*27*). We used a zero-phase forward reverse digital filtering to avoid any spurious time delay between the vocalization and EMG signals due to bandpass filtering. The vocalization was simultaneously recorded using Plexon® omniplex. We detected the onset and offset of the syllable production in the same way as in the undirected experiments. The numbers of cry syllables for each infant were (n = 378, 385, 228, 364, 457) and of phee syllables were (n = 288, 239, 130, 149, 202).

Dynamic time warping (DTW) analysis

To measure the similarity between two time series (respiratory EMG signals) with possibly different time lengths, we used the continuous DTW algorithm using a linear interpolation model. We used the routine implemented by Pau Micó (MATLAB® file exchange #16350). The cost of the DTW was used as a measure of similarity between two signals. Smaller values of DTW cost indicate larger similarity between the signals. We calculated the DTW cost between all pairs of cry EMG and for all pairs of phee EMG for each infant. We used the two-way ANOVA to compare the mean EMG similarities (DTW cost) between cries and phees for each infant. The two-way ANOVA was used to control for the effect of variability between individuals. The post-hoc analysis was done with Bonferroni correction.

Calculation of phee-cry ratio and zero-crossing day

For the directed calls experiments, we defined as a whole (i.e., multisyllabic) call as any uninterrupted sequence of utterances of the same type (phee or cry) separated (previous offset to next onset) by less than 500ms (*8, 28*). To quantify the developmental transition from cry to phee, for each session and subject, we calculated the ratio between the number of phee minus cry and the number of phee plus cry, *i.e.*,

$$\text{phee/cry ratio} = \frac{\#\ of\ infant\ phee\ calls\ produced - \#\ of\ infant\ cry\ calls\ produced}{\#\ of\ infant\ phee\ calls\ produced + \#\ of\ infant\ cry\ calls\ produced}.$$

A phee/cry ratio that is greater than zero corresponds to a larger production of phee in comparison to cry, while a phee/cry ratio less than zero corresponds to more production of cry in relation to phee. To represent the change in phee/cry ratio across

development, we fitted a cubic spline curve to the data and the resulting curve was called *phee/cry ratio curve*. We called *zero- crossing day* the first point at which the phee/cry ratio curve was equal to zero, transitioning from a negative phee/cry ratio to a positive phee/cry ratio. The idea of the zero-crossing day is that it quantifies how fast each infant transition from cry abundant initial period to phee dominated later period. We tested if the rate of weight change before the zero-crossing day could predict the change in phee/cry ratio. For this, we first calculated the difference between two consecutive weight measurements and divided by the number of days between the measurements. This gives us the local rate of weight change. The rate of weight change was calculated as the average of local rate of weight changes before the zero-crossing day. If there were any monotonic relationship between the weight change and the timing of transition from cries to phees, we would expect a significant Spearman correlation (r) between the rate of weight change and the zero-crossing day. We also calculated the robust linear regression curve and the two-sided t-test of the nullity of regression slope (n = 10 infants).

Contingent/non-contingent vs zero crossing day

A parental call was classified as contingent response to an infant call if the onset of parental call was separated by less than 5s from the offset of the infant call and there is no other call between both calls (*8, 29*). To test if the contingent parental responses were related with how fast the infants transition from cry to phee, we calculated the correlation between the proportion of infant phees for which the parents responded before the zero-crossing day (total number of contingent parental responses before the zero-crossing day divided by the total number of infant phees in the period) and the zero-crossing day. To calculate the correlation, we included only the proportion of contingent parental responses that happened before the zero-crossing day to be consistent with the causal ordering where the possible cause (contingent parental response) happens before the effect (zero-crossing day). We calculated the robust linear regression curve and used a two-sided t-test to verify if the slope of the regression was significantly different from zero (n = 10 infants). We also calculated the Spearman correlation to measure the strength of interaction. As a control for the contingent response calls, we tested if the non-contingent parental calls were related with how fast the infants transition from cry to phee. To do this, we calculated the Spearman correlation between the proportion of parental phees that were not contingent before the zero-crossing day (total number of non-contingent parental phees before the zero-crossing day divided by the total number of parental phees in the period) and the zero-crossing day. We also calculated the robust linear regression curve and the two-sided t-test of the nullity of regression slope. To account for the possibility that the observed correlation between the proportion of infant phees for which the parents responded and the zero-crossing-day is a consequence of genetics, we calculated, for six infants that had identical parents, the Spearman correlation between the proportion of infant phees for which the parents responded before the zero-crossing day and the zero-crossing-day. Because of small sample size, to test the nullity of Spearman correlation, we used the exact permutation test, splitting the proportion of ties. We also tested if the coefficients of the robust linear regression using the data for all infants and for the six infants with same parents were statistically different. For this, we applied a bootstrap test, in which we randomly resampled with replacement six infants and calculated the coefficients of the robust linear fit between the

proportion of phees for which parents responded and the zero-crossing day. We repeated this procedure 10,000 times and calculated the p-value under the null hypothesis of equality of coefficients.

To verify if there was any significant change in parental vocal output during infant development, we calculated the rate of parental phee call production. For each parent (mother and father), we calculated the rate of phee production during vocal interaction with each infant. We fitted for each parent-infant pair a cubic spline curve (MATLAB® csap) to represent the trajectory of phee production during infant development. We also fitted a cubic spline curve to the population data of phee production for the mother and father. To verify if there is any systematic change in the phee production rate during infant development, we fitted a robust linear regression and used the two-sided t-test of the nullity of regression slope (n = 98 sessions for mother and n = 97 sessions for father).

Supplementary Text

Growth is not linearly related to changes in acoustic parameters

In the main text, we showed that the overall growth measured by weight is not linearly correlated with the developmental change of acoustic parameters. It is still possible that a higher order polynomial relationship could capture better the relationship between the acoustic parameters and weight. A standard way to infer possible non-linear relationship is to use log transforms of the variables. In this way, if the variables are related by some polynomial equation (e.g., $y = a*x^p$), the log transform will linearize the relationship (e.g., $\log(y) = p \log(x) + \log(a)$) and then standard linear statistical inference can be applied. Fig. S1B shows that log weight changes visibly contrast with trajectories of the log transformed acoustic parameters (fig. S1A). To quantify this difference, we used log weight to predict changes in the four log transformed acoustic parameters. Black curves in fig. S1A represent predicted average parameter values given the average weight for each postnatal day and fig. S1C shows residues of these predictions. If growth completely explained the acoustic parameters, the residues should be uncorrelated and identically distributed across postnatal days. Using Akaike Information Criterion (AIC), the order of the best polynomial fit was 3 for all residues related to the four acoustic parameters (Fig. 1D, main text). These results are very similar to the results without the log transforms and the conclusion is exactly the same. To show this, we calculated the correlation between the residues for original and log transformed variables, respectively, for duration, dominant frequency, amplitude modulation frequency, and Wiener entropy (Spearman correlation = 0.943, 0.974, 0.963, -0.956; $p < 0.001$ for all four parameters).

Fig. S1

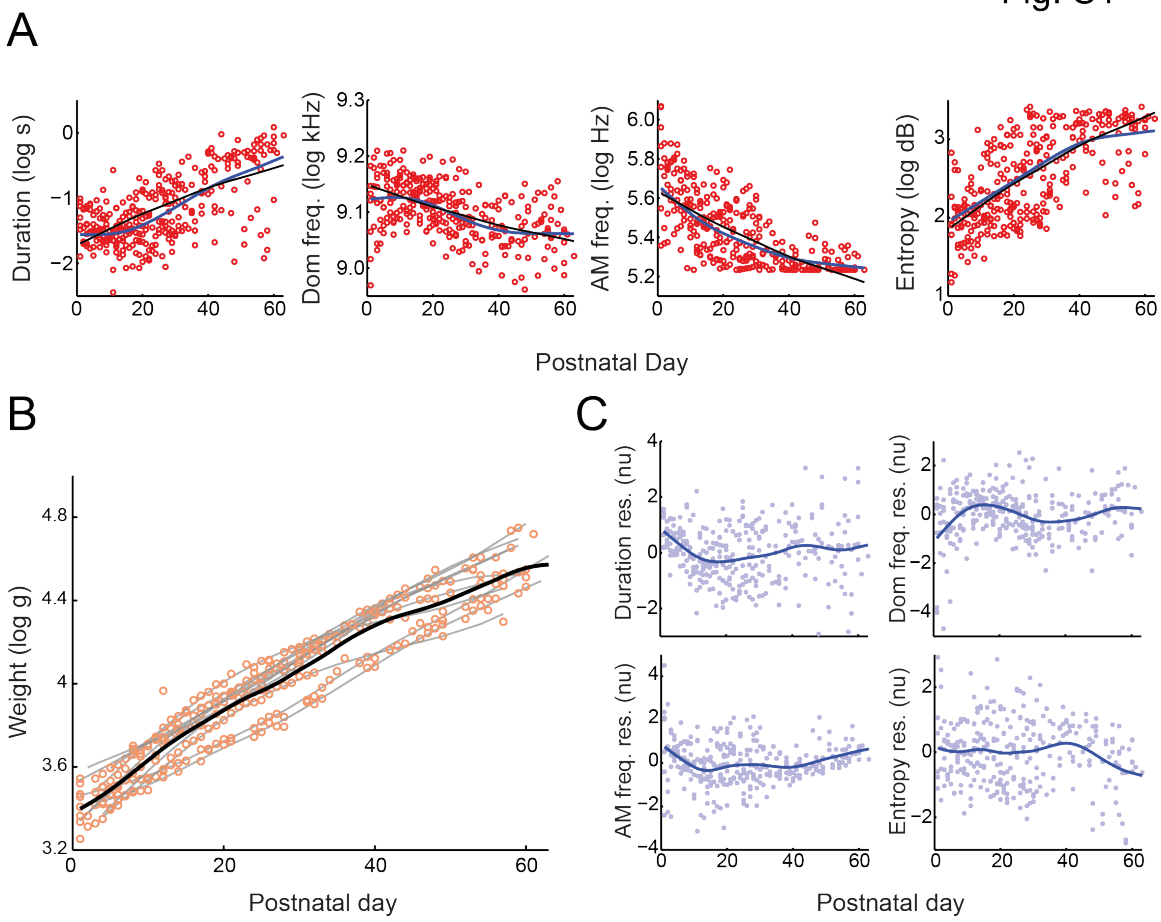

Fig. S1. Log weight cannot predict the developmental change in the log transformed acoustic parameters. (A) Scatter plots of developmental changes of four log transformed acoustic parameters for all 10 infants, showing average values per session for each infant (red circles) and cubic spline fit for the population average (blue curves). Black curves show log transformed parameter values through development predicted by the animals' average daily weights. (B) Individual log transformed weights of each infant during development (orange circles) and cubic spline fit (gray curves) for their weight changes; black curve is cubic spline fit for average population log transformed weight. (C) Standardized residues of the regression using log transformed weights as predictors (blue points) and cubic spline fitted to residues in normalized units (blue curves).

Fig. S2

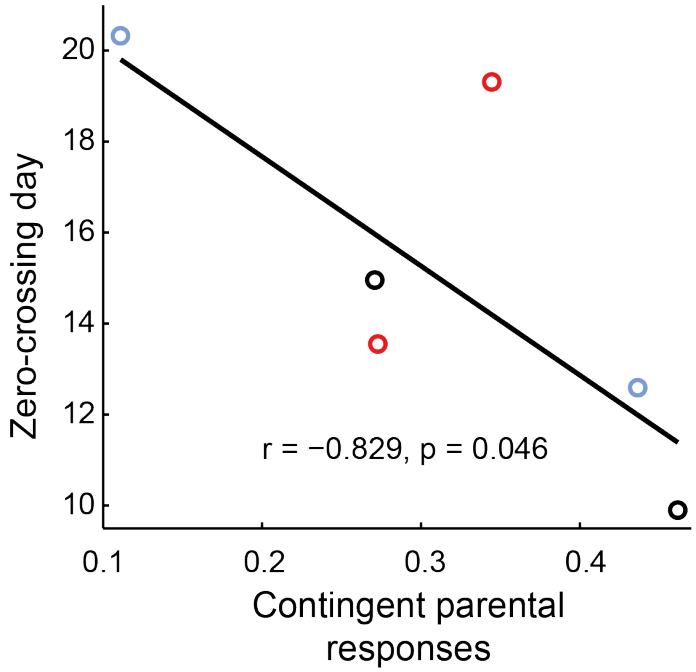

Fig. S2. Proportion of contingent parental calls and zero-crossing day for full-siblings are correlated. The circles indicate the proportion of contingent parental response and the respective zero-crossing day. All six infants in this figure have identical parents. Circles with same colors indicate dizygotic twins. The black line represents the robust linear regression fit.

References (22-29)

22. von Luxburg, U. A tutorial on spectral clustering. Statistics and Computing **17**, 395-416 (2007).

23. Fujita, A., Takahashi, D. Y. & Patriota, A. G. A non-parametric method to estimate the number of clusters. Computational Statistics & Data Analysis **73**, 27-39 (2014).

24. Titze, I. R. The physics of small-amplitude oscillation of the vocal folds. The Journal of the Acoustical Society of America **83**, 1536-1552 (1988).

25. Perl, Y. S., Arneodo, E. M., Amador, A., Goller, F. & Mindlin, G. B. Reconstruction of physiological instructions from zebra finch song. Physical Review E **84**, 051909 (2011).

26. Kasdin, J. Discrete simulation of colored noise and stochastic processes and 1/fa power law noise generation. Proceedings of the IEEE **83**, 802-827 (1995).

27. Clancy, E. A. Electromyogram amplitude estimation with adaptive smoothing window length. IEEE Transactions on Biomedical Engineering **46**, 717-729 (1999).

28. DiMattina, C. & Wang, X. Virtual vocalization stimuli for investigating neural representations of species-specific vocalizations. Journal of Neurophysiology **95**, 1244-1262 (2006).

29. Miller, C. T. & Wang, X. Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. Journal of comparative physiology A **192**, 27-38 (2006).

Audio S1 – S8

Sound files corresponding to the spectrograms in Fig. 1A which illustrate the dramatic changes in vocal output of infant marmosets over the course of 2 months in the undirected (social isolation) context. Audio S1 to S4 correspond to the spectrograms for Infant 1 at P1, P13, P36, and P60, respectively. Audio S5 to S8 correspond to the spectrograms for Infant 2 at P1, P14, P33, and P59, respectively.

Supplementary data

Data set used to plot the figures and calculate the statistics in the article. SupplementaryData.zip file contains six .mat files that we describe below.

1) AcousticParameter_Development.mat (Figs. 1B, 1D, 2A, 2B, 2C)

- ParameterValuePerSubjDay(Subject, Day, Params) gives the average acoustic parameter value of Params (1 = duration, 2 = dominant frequency, 3 = AM frequency, 4 = Wiener entropy) for each Subject (marmoset 1 to 10) at Day (postnatal day 1 to 63). It returns NaN if there was no recording for that Day and Subject.
- ParameterValuePerSubjDaySyllable{Subject}{Session}(Syllable, Params) gives the parameter value of Params (1 to 4) for each Syllable and Subject (1 to 10) at Session. The corresponding dates for the Session are given by the postnatal days that are not NaN in ParameterValuePerSubjDay.
- ParameterName{Params} gives the name of parameters corresponding to Params.

2) Weights.mat (Figs. 1C, 1D)

- Weightdata{Subject}(Session, [Day Weight]) gives for each Subject (1 to 10) and Session the corresponding Day (postnatal day) and Weight.

3) CallDuration_Development.mat (Fig. 2D)

- TotDurationCall{Subject}{Type}(Session) gives the total duration (sum of the duration of all calls in a session) for each Session and given Type (1= Phee, 2 = Twitter, 3 = Trill, 4 = Cry, 5 = Subharmonic-phee, 6 = Phee-cry).
- CallName{Type} gives the name of the call type for each Type.

4) DTW.mat (Fig. 3K)

- DTWdist{Type}{Subject} returns the DTW costs for all the pair of syllables for each Type (1 = Phee, 2 = Cry) and Subject (1 to 5).
- CallType{Type} gives the name of the call type for each Type.

5) PheeCryTransition.mat (Figs. 4A, 4B, 4C, 4D, 4E, 4F)

- ContingentParentalResp(Subject) gives the proportion of infant calls for which the parents responded before the zero-crossing day for each Subject (1 to 10)

- NonContingentParentalResp(Subject) gives the proportion of non-contingent parental calls before the zero-crossing day for each Subject (1 to 10).
- Cry{Subject}(Session) gives the total number of cries produced in Session for each Subject (1 to 10).
- Phee{Subject}(Session) gives the total number of phees produced in Session for each Subject (1 to 10).
- PheeCryRatio{Subject}(Session) gives the phees-cry ratio for each Session and Subject (1 to 10).
- PostnatalDay{Subject}(Session) gives the corresponding postnatal day for each Session and Subject (1 to 10).
- RateWeightChange(Subject) gives the rate of weight change for each Subject (1 to 10).
- ZeroCrossingDay(Subject) gives the zero-crossing day for each Subject (1 to 10).

6) MotherFatherCallRate.mat (Fig. 4G)

- MotherPheeCallRate(Session) gives the number of phees produced per minute by the mother in a Session (1 to 98).
- MotherPND(Session) gives the postnatal day of each Session for the mother (1 to 98).
- FatherPheeCallRate(Session) gives the number of phees produced per minute by the father in a Session (1 to 97).
- FatherPND(Session) gives the postnatal day of each Session for the father (1 to 97).