

The autonomic nervous system is the engine for vocal development through social feedback

Asif A Ghazanfar^{1,2} and Yisi S Zhang^{1,2}



At least one non-human primate species — the marmoset monkey — exhibits developmental processes similar to human vocal development. These processes include babbling-like early vocal output and a role for social feedback in changing this output into mature-sounding vocalizations. Such parallel behaviors provide a window through which we can begin to understand the physiological mechanisms for how early vocalizations are produced and shaped by social feedback. The latest work shows that the acoustic structure of babbling in infant monkeys is driven by oscillations of the autonomic nervous system. It is hypothesized that this autonomic nervous system rhythm is perturbed through vocal interactions between infants and parents. These interactions gradually accelerate the transformation of immature vocalizations into mature forms.

Addresses

¹ Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA

² Department of Psychology, Princeton University, Princeton, NJ 08544, USA

Corresponding author: Ghazanfar, Asif A (asifg@princeton.edu)

Current Opinion in Neurobiology 2016, **40**:155–160

This review comes from a themed issue on **Systems neuroscience**

Edited by **Don Katz** and **Leslie Kay**

<http://dx.doi.org/10.1016/j.conb.2016.07.016>

0959-4388/© 2016 Elsevier Ltd. All rights reserved.

Introduction

Like many behaviors [1^{**}], vocal development is the outcome of interactions between the infant's developing biological system of vocal production (the body and the nervous system) and his experience with caregivers. In humans, such development is marked by a progression: early vocalizations, like cries, fussing and cooing sounds, become increasingly complex and speech-like and eventually turn into words [2]. Early vocalizations act as scaffolding for later, more speech-like vocalizations. In fact, from a purely acoustic perspective, the progression of changes from cries to later vocalizations like cooing and babbling is continuous [3].

Social feedback from caregivers can be an instrumental driving force that increases the maturation rate of these

prelinguistic vocalizations [4]. Not only is the volubility of infants influenced by social context and past dyadic interactions with caregivers [5,6], but caregivers who preferentially and contingently respond to speech-like infant vocalizations spur the development of more complex vocalizations from those infants [7^{**},8^{**}]. Conversely, parents are more likely to contingently respond to infant vocalizations if they sound more mature or speech-like [9]. Taken together, these findings show that contingent parental responses to infants' vocalizations facilitate the latter's development into more mature sounding forms.

In order to understand the mechanisms underlying these vocal development processes, we need an animal model system that shares at least some of these features of human infant vocal development and parent–infant interactions. Below, we review what we have learned from marmoset monkey vocal development and show they that have a similar developmental trajectory: Marmosets produce babbling-like vocalizations and the maturation rate of their vocalizations is influenced by parental feedback. We will then present evidence for a physiological mechanism that guides this developmental process.

The marmoset monkey model system

The marmoset monkey (*Callithrix jacchus*) has recently come into the fore as a new primate model system for neuroscience and communication [10,11]. Marmosets and other closely related species in the Callitrichid family are quite flexible in their vocal behavior, especially when compared to Old World primates like macaques and apes. They readily adjust (without any training) the timing of their vocalizations to the timing of conspecific calls [12–15], and they cooperatively adjust the amplitude of their calls during vocal exchanges in accord with distance from conspecifics [16]. Marmosets also take turns when they vocalize, exhibiting contingent and repeated exchanges of vocalizations between any two individuals — related or unrelated — for an extended period of time [15]. This turn-taking behavior by marmosets has the same universal features, and coupled oscillator properties, of human conversational turn-taking (albeit on a different time-scale) [15,17,18].

Marmoset monkeys are also cooperative breeders. Both parents, older siblings and non-kin help care for offspring by carrying them and sharing food with them. Among primates, this is very rare: only humans and members of the Callitrichid family exhibit this reproductive strategy.

Thus, in terms of comparative developmental studies among human and nonhuman primates, marmosets are a more compelling analogous species than the phylogenetically closer, but socially dissimilar, Old World apes and monkeys [19]. These cooperative breeding behaviors by humans and marmosets are the basis for other prosocial cognitive processes [20,21], including those related to vocal communication [11]. In humans, the vocal behavior of infants stimulates interactions with caregivers, thereby serving a key role in the infants' own development. Thus, in light of their similar infant care strategies and vocal turn-taking behavior, does the vocal behavior of marmoset monkeys develop in a manner similar to that of human vocal behavior?

Marmoset monkey vocal development is influenced by parental feedback

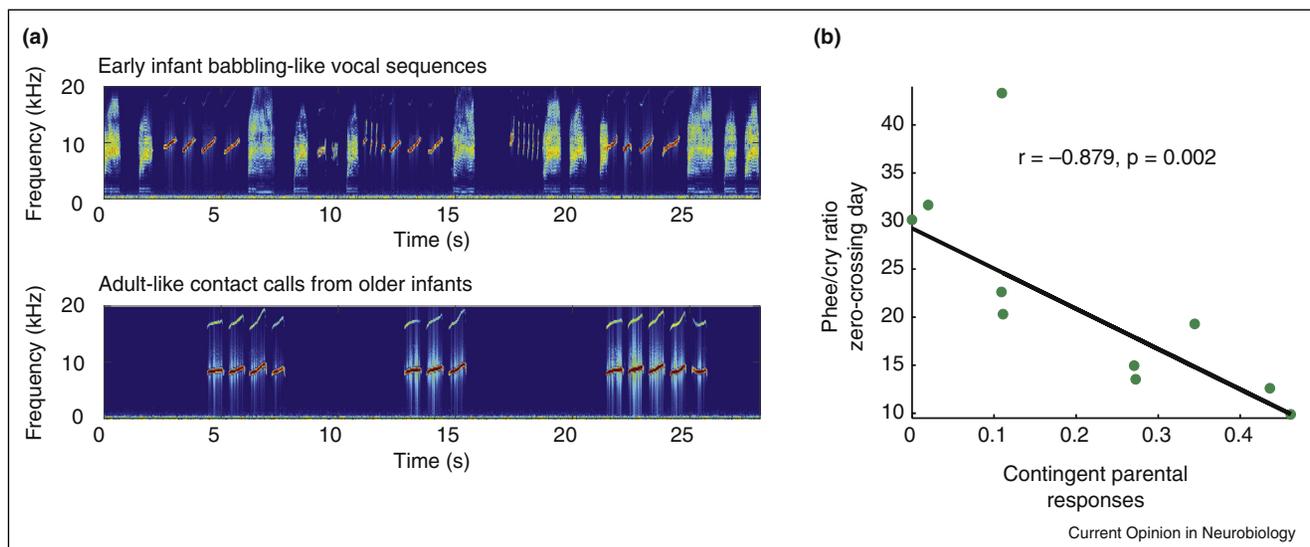
Human infant babbling is a mixture of vocalizations — a subset of which will be incorporated in adult speech and produced in correct contexts, while others are transient, serving as substrate for later vocalizations [2]. Such transient vocalizations are also evident during song learning by birds [22**]. Unlike Old World nonhuman primates, developing marmosets also produce a babbling-like vocal output [23,24**,25], with some vocalizations that sound adult-like (but are produced in the wrong social context) and others that are immature forms of what will ultimately be the contact ('phee') call used by them during vocal exchanges (Figure 1a) [24**,26**].

It is typically assumed that any developmental changes in nonhuman primate vocalizations are simply related to growth [27]. That is, any changes in vocal structure are just simple consequences of biomechanical and/or neural circuit maturation and that social feedback from parents has no influence on these changes. Marmoset monkeys, however, are unlike the typical primate. The rate at which their vocalizations (specifically, their contact 'phee' call) matures is strongly influenced by how much contingent vocal feedback they get from their parents [24**]. Figure 1b shows that the timing of the transition point between producing more mature phee calls versus immature cries is significantly correlated with how often parents contingently respond to infant vocalizations. By contrast, there was no relationship between the rate of phee call maturation and the total amount of parental phee calls heard [24**]. Taken together, these findings provided an opportunity to study the physiological mechanism underlying human-like vocal development in a primate species.

Rhythmicity in babbling: an arousal mechanism

As it is for human infants [28] and songbirds [22**,29], the babbling output of marmoset infants is very rhythmic [26**]. This rhythmicity suggests that this output is driven by the oscillatory activity of the nervous system. As in developing songbirds [30], the temporal structure of babbling output by infants is tightly locked to respiration

Figure 1



Marmoset infant vocalizations are influenced by contingent vocal responses from parents. (a) Spontaneous babbling-like vocalizations from an infant marmoset in the first week of life (postnatal day 3; top panel) and later in infancy (postnatal day 40; bottom panel). These vocalizations were recorded in an undirected (social isolation) context. The early vocalizations are highly variable whereas later in development, infants only produce contact ('phee') calls in this undirected context. (b) The more contingent vocal responses an infant hears following its own vocalizations, the faster it will transform high entropy calls (cries) into low entropy (phee) calls. Zero-crossing day is an index of vocal development, referring to the postnatal day on which an infant produces calls with a 50–50 phee/cry ratio.

in marmoset monkey infants [26^{••}], with respiration providing the power for generating vocal sounds [31]. Approximately every second, an utterance phase-locked to the respiratory rhythm is produced. The time-varying spectral structure of babbling sequences also has a rhythm but at a rate that is an order of magnitude slower. Spectral entropy, a measure of the noisiness of the sound spectrum [24^{••},32], fluctuates during babbling at a 0.1 Hz frequency [26^{••}]. This is interesting because the autonomic nervous system (ANS) also has a 0.1 Hz rhythm known as the Mayer wave. Thus, in effect, arousal is oscillating at this frequency.

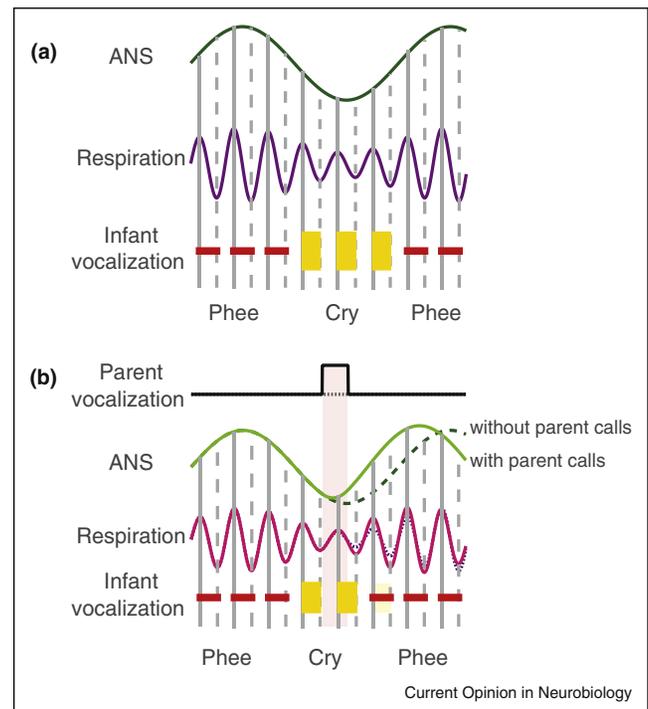
Present in all mammals, the Mayer wave represents perturbations to the baroreflex [33], the homeostatic mechanism by which blood pressure is maintained via changes in heart rate. This results in an unstable negative feedback control loop that generates self-sustained oscillations at its resonance frequency of 0.1 Hz. How can this ANS rhythm account for the 0.1 Hz entropy fluctuations in infant babbling sequences? The 0.1 Hz ANS rhythm modulates the power of the faster, 1 Hz respiratory rate [26^{••}] (Figure 2a). This creates a scenario in which marmoset infants produce vocalizations at a rate of 1 Hz but the respiratory power is not constant — it is modulated according to a 0.1 Hz rhythm. Lower respiratory power generates noisy (high entropy) vocalizations, while higher respiratory power generates more tonal (low entropy) vocalizations [24^{••},26^{••}] (Figure 2a). Thus, as revealed by partial coherence measures of heart rate, respiration and vocalizations, spectral entropy fluctuates at ~ 0.1 Hz because both laryngeal tension and respiratory power are modulated by the ANS rhythm [26^{••}].

Although the rhythmicity of marmoset infant babbling is explained by this ANS rhythm, three other features of marmoset vocal development have yet to be accounted for mechanistically: (1) infant marmosets gradually stop producing immature vocalizations and start producing only mature phee calls in the undirected (isolation) (Figure 1a) and directed (vocal interactions with out-of-sight conspecifics) contexts [24^{••},26^{••}]; (2) the transformation of immature vocalizations into phee calls is influenced by the amount of contingent parental vocal responses [24^{••}] (Figure 1b); and (3) phee calls by adults are spontaneously produced at ~ 0.1 Hz frequency [34].

Raising the threshold to vocalize: a hypothesis

A simple developmental process may be at play for all three changes, one that involves raising the ‘decision’ threshold relative to the ANS rhythm. It is our hypothesis that contingent vocal feedback from parents can accelerate the rising of this threshold and thus the production of mature sounding contact calls at the adult rate of production (Figure 2b). Figure 3a illustrates how this would work. In this scheme, very young infant marmosets have a very low threshold to vocalize, such that they produce

Figure 2

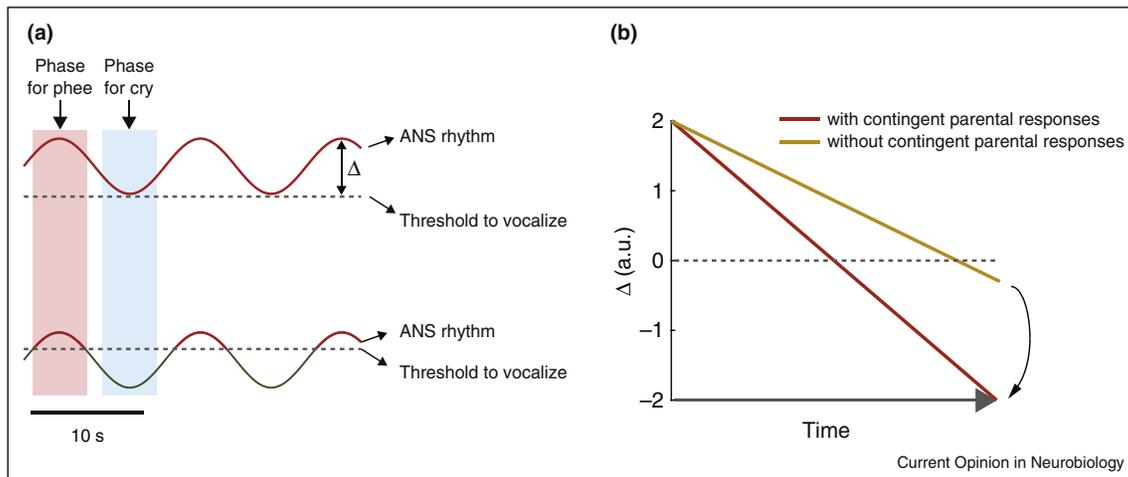


The autonomic nervous system modulates arousal in a rhythmic fashion that in turn modulates respiratory power. (a) A schematic illustrating spontaneous vocal production as a function of ANS oscillation. The continuously produced vocalizations by very young infant marmosets are driven by the natural rhythmic activity of respiration whose power is modulated by the slower, ~ 0.1 Hz rhythm of the ANS. This consequently changes the quality of the vocalizations so that they fluctuate between high (cry) and low (phee) levels of entropy. (b) Changes in infant vocalizations might result from the perturbation of its ANS rhythm upon hearing a parental vocal response. The schematic shows that when hearing a parental call, the phase of the infant's ANS oscillation shifts towards the peak, resulting in an enhancement of respiratory amplitude and thus a lower entropy, more mature sounding, vocalization.

vocalizations at the respiration rate of ~ 1.0 Hz (but with spectral entropy changing at 0.1 Hz due to the ANS modulation of respiratory power; see above). Over the course of development, this threshold rises to the point where only the peaks of the ANS rhythm exceed it. This results in two outcomes: only phee calls are produced and they are produced at the adult rate of ~ 0.1 Hz.

In the framework of ANS-governed infant vocal production, the effect of caregiver's contingent responses might be to provide a transient driving force to push the ANS oscillator towards the peak so that the oscillatory phase corresponding to producing adult-like vocalizations emerges immediately after the parental call (Figure 2b). The indirect evidence for this process consists of the following: First, after a contingent vocal response from a parent, marmoset infants immediately produce more mature-sounding vocalizations [35^{••}].

Figure 3



A simple threshold model of vocal development that can account for the transition from babbling to mature vocalizations. **(a)** Different phases of the oscillation correspond to different levels of spectral entropy in the vocalizations that are produced: low entropy phee calls and high entropy cries. A newborn's ANS oscillates above the threshold, allowing continuous production of various calls at the faster respiration rate. As the infant develops, the threshold is elevated and only phases close to the peak can be above the threshold. This results in the production of only phee calls at every ANS cycle. **(b)** The rate of changes in the proportion of phee calls might be related to the rate the threshold is elevated through development, and this rate is influenced by contingent parental vocal responses.

For example, the ANS-dependent spectral entropy of their vocalizations was lower in the 5-second interval following a contingent parental vocal response than in the 5 s preceding it. Thus, marmoset infants not only change the structure of their vocalizations over the course of days, but do so in real time as they interact with their parents. In humans, contingent caregiver vocal responses to infant vocalizations also immediately result in more mature, speech-like vocalizations from those infants [8^{**}]. Second, marmoset parents are more likely to give contingent responses to infant vocalizations that sound more mature (i.e., have lower entropy) [35^{**}]. This, too, is like human infant-caregiver interactions [9]. These data show that, in both species, early vocal learning by infants is a system property that includes both the infant and parents as important elements [35^{**}].

Although such perturbations to the infant's ANS/vocal production system is brief, cumulatively, it might change the infant vocal behavior in the long term and this process might account for the elevation of the hypothesized threshold relative to ANS rhythm (Figure 3a,b).

Conclusions

Vocal signals are part of a complex, multidimensional, probabilistic process that includes the physiological states of all the participants [36]. Nowhere is this more important than during caregiver-infant interactions and their influence on infant vocal development. From a translational perspective, neurodevelopmental disorders representing two opposite extremes in social communication — autism and Williams syndrome — have been linked to arousal/ANS

dysfunction [37,38]. Thus, in order to understand early vocal development (and how it may go awry), it is critical to know how the ANS may function in producing individual differences in vocal output.

These developmental data from marmoset monkeys support a cybernetic-like conceptual framework (similar but not identical to the one put forth by Zeskind [36]) wherein there are four basic elements to infant-caregiver vocal interactions: (1) early vocalizations of infants are largely governed by the rhythmical dynamics of the ANS/arousal; (2) arousal dynamics are reflected in the acoustic structure of infant vocalizations; (3) these differences in the acoustic structure of infant vocalizations influence how caregivers respond; and (4) the contingent responses of caregivers subsequently influence the acoustic structure of infant vocalizations. Thus, using marmoset monkeys as a model system is revealing how rhythmic fluctuations of the ANS may be of critical importance to understanding the early vocal development in humans and other species, acting as the engine for vocal development through social interactions.

Conflicts of interest statement

Nothing declared.

Acknowledgements

We thank Jeremy Borjon, Diana Liao and Daniel Takahashi for their influence in shaping many of the ideas presented in this manuscript. The National Institutes of Health (NINDS) R01NS054898 (AAG) and a James S. McDonnell Foundation Scholar Award (AAG) supported this work

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Byrge L, Sporns O, Smith LB: **Developmental process emerges from extended brain–body–behavior networks.** *Trends Cogn Sci* 2014, **18**:395–403.

An outstanding essay emphasizing that any time of behavioral development can only be understood by investigating the interplay of brain–body–behavior networks, spanning timescales. We took this perspective in the current review.

2. Oller DK: *The Emergence of the Speech Capacity*. Lawrence Erlbaum Associates, Inc.; 2000.
3. Kent RD, Murray AD: **Acoustic features of infant vocalic utterances at 3, 6, and 9 months.** *J Acoust Soc Am* 1982, **72**:353–365.
4. Gros-Louis J, West MJ, King AP: **Comparative perspectives on the missing link: communicative pragmatics.** In *The Oxford Handbook of Developmental Behavioral Neuroscience*. Edited by Blumberg MS, Freeman JH, Robinson SR. Oxford University Press; 2010:684–707.
5. Franklin B, Warlaumont AS, Messinger D, Bene E, Nathani Iyer S, Lee C-C, Lambert B, Oller DK: **Effects of parental interaction on infant vocalization rate, variability and vocal type.** *Lang Learn Dev* 2013, **10**:279–296.
6. Goldstein MH, Bornstein MH, Schwade JA, Baldwin F, Brandstadter R: **The value of vocalizing: five-month-old infants associate their own non-cry vocalizations with responses from caregivers.** *Child Dev* 2009, **80**:636–644.
7. Goldstein MH, King AP, West MJ: **Social interaction shapes babbling: testing parallels between birdsong and speech.** *Proc Natl Acad Sci U S A* 2003, **100**:8030–8035.

Refs. [7**,8**] are two important, well-controlled studies demonstrating the influence of social feedback on human infant vocal development. They draw parallels to important but often ignored processes of vocal learning in songbirds that are not related to imitation.

8. Goldstein MH, Schwade JA: **Social feedback to infants' babbling facilitates rapid phonological learning.** *Psychol Sci* 2008, **19**:515–523.
- See annotation to Ref. [7**].
9. Gros-Louis J, West MJ, Goldstein MH, King AP: **Mothers provide differential feedback to infants' prelinguistic sounds.** *Int J Behav Dev* 2006, **30**:509–516.
 10. Miller CT, Freiwald WA, Leopold DA, Mitchell JF, Silva AC, Wang X: **Marmosets: a neuroscientific model of human social behavior.** *Neuron* 2016, **90**:219–233.
 11. Borjon JI, Ghazanfar AA: **Convergent evolution of vocal cooperation without convergent evolution of brain size.** *Brain Behav Evol* 2014, **84**:93–102.
 12. Ghazanfar AA, Flombaum JI, Miller CT, Hauser MD: **The units of perception in the antiphonal calling behavior of cotton-top tamarins (*Saguinus oedipus*): playback experiments with long calls.** *J Comp Physiol A Sens Neural Behav Physiol* 2001, **187**:27–35.
 13. Ghazanfar AA, Smith-Rohrberg D, Pollen AA, Hauser MD: **Temporal cues in the antiphonal long-calling behaviour of cottontop tamarins.** *Anim Behav* 2002, **64**:427–438.
 14. Miller CT, Beck K, Meade B, Wang X: **Antiphonal call timing in marmosets is behaviorally significant: interactive playback experiments.** *J Comp Physiol A* 2009, **195**:783–789.
 15. Takahashi DY, Narayanan DZ, Ghazanfar AA: **Coupled oscillator dynamics of vocal turn-taking in monkeys.** *Curr Biol CB* 2013, **23**:2162–2168.
 16. Choi JY, Takahashi DY, Ghazanfar AA: **Cooperative vocal control in marmoset monkeys via vocal feedback.** *J Neurophysiol* 2015, **114**:274–283.
 17. Levinson SC: **Turn-taking in human communication — origins and implications for language processing.** *Trends Cogn Sci* 2016, **20**:6–14.

18. Ghazanfar AA, Takahashi DY: **The evolution of speech: vision, rhythm, cooperation.** *Trends Cogn Sci* 2014, **18**:543–553.
19. Elowson AM, Snowdon CT, Lazaro-Perea C: **'Babbling' and social context in infant monkeys: parallels to human infants.** *Trends Cogn Sci* 1998, **2**:31–37.
20. Burkart JM, Allon O, Amici F, Fichtel C, Finkenwirth C, Heschl A, Huber J, Isler K, Kosonen ZK, Martins E *et al.*: **The evolutionary origin of human hyper-cooperation.** *Nat Commun* 2014, **5**:4747.
21. Snowdon CT, Cronin KA: **Cooperative breeders do cooperate.** *Behav Processes* 2007, **76**:138–141.
22. Tchernichovski O, Mitra PP, Lints T, Nottebohm F: **Dynamics of the vocal imitation process: how a zebra finch learns its song.** *Science* 2001, **291**:2564–2569.

A pioneering study of the trajectories of vocal change during song learning by zebra finches. It quantitatively showed a number of unexpected processes including non-monotonic changes in song syllables and the differentiation of syllables from a prototype.

23. Elowson AM, Snowdon CT, Lazaro-Perea C: **Infant 'babbling' in a non-human primate: complex vocal sequences with repeated call types.** *Behaviour* 1998, **135**:643–664.
 24. Takahashi DY, Fenley AR, Teramoto Y, Narayanan DZ, Borjon JI, Holmes P, Ghazanfar AA: **The developmental dynamics of marmoset monkey vocal production.** *Science* 2015, **349**:734–738.
- The very first nonhuman primate study to quantitatively show dramatic changes in vocalizations during development and that the maturation of vocalizations are influenced by contingent parental vocal feedback.

25. Pistorio AL, Vintch B, Wang X: **Acoustic analysis of vocal development in a New World primate, the common marmoset (*Callithrix jacchus*).** *J Acoust Soc Am* 2006, **120**:1655–1670.
 26. Zhang YS, Ghazanfar AA: **Perinatally influenced autonomic nervous system fluctuations drive infant vocal sequences.** *Curr Biol* 2016, **26**:1249–1260.
- An electromyography study of the vocal babbling in infant marmosets that linked the rhythmic fluctuations of the autonomic nervous system to changes in respiratory power and thus to changes in the spectral and temporal properties of vocal sequences.
27. Egnor SER, Hauser MD: **A paradox in the evolution of primate vocal learning.** *Trends Neurosci* 2004, **27**:649–654.
 28. MacNeilage PF: *The Origin of Speech*. Oxford University Press; 2008.
 29. Sasahara K, Tchernichovski O, Takahashi M, Suzuki K, Okanoya K: **A rhythmic landscape approach to the developmental dynamics of birdsong.** *J R Soc Interface* 2015, **12**:20150802.
 30. Veit L, Aronov D, Fee MS: **Learning to breathe and sing: development of respiratory-vocal coordination in young songbirds.** *J Neurophysiol* 2011, **106**:1747–1765.
 31. Ghazanfar AA, Rendall D: **Evolution of human vocal production.** *Curr Biol* 2008, **18**:R457–R460.
 32. Tchernichovski O, Nottebohm F, Ho CE, Pesaran B, Mitra PP: **A procedure for an automated measurement of song similarity.** *Anim Behav* 2000, **59**:1167–1176.
 33. Julien C: **The enigma of Mayer waves: facts and models.** *Cardiovasc Res* 2006, **70**:12–21.
 34. Borjon JI, Takahashi DY, Cervantes DC, Ghazanfar AA: **Arousal dynamics drive vocal production in marmoset monkeys.** *J Neurophysiol* 2016. in press.
 35. Takahashi DY, Fenley AR, Ghazanfar AA: **Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys.** *Phil Trans R Soc B* 2016, **371**:20150370.
- This study showed that infant marmosets immediately produce more mature sounding vocalizations following contingent parental vocal responses and that parents are more likely to provide vocal feedback if the infant's call sounds more adult-like. These processes are exactly the same as what has been reported from human infant–caregiver vocal interactions.
36. Zeskind PS: **Infant crying and the synchrony of arousal.** In *The Evolution of Emotional Communication: From Sounds in*

Nonhuman Mammals to Speech and Music in Man. Edited by Altenmüller E, Schmidt S, Zimmermann E. Oxford University Press; 2013:155-174.

37. Bal E, Harden E, Lamb D, Van Hecke AV, Denver JW, Porges SW: **Emotion recognition in children with autism spectrum disorders: relations to eye gaze and autonomic state**. *J Autism Dev Disorders* 2010, **40**:358-370.
38. Jarvinen AM, Bellugi U: **What does Williams syndrome reveal about the determinants of social behavior?** *Front Hum Neurosci* 2013, **7**:321.