

Constraints and flexibility during vocal development: insights from marmoset monkeys

Asif A Ghazanfar^{1,2,3} and Diana A Liao¹



Human vocal development is typically conceived as a sequence of two processes — an early maturation phase where vocal sounds change as a function of body growth ('constraints') followed by a period during which social experience can influence vocal sound production ('flexibility'). However, studies of other behaviors (e.g., locomotion) reveal that growth and experience are interactive throughout development. As it turns out, vocal development is not exceptional; it is also the on-going result of the interplay between an infant's growing biological system of production (the body and the nervous system) and experience with caregivers. Here, we review work on developing marmoset monkeys — a species that exhibits strikingly similar vocal developmental processes to those of prelinguistic human infants — that demonstrates how constraints and flexibility are parallel and interactive processes.

Addresses

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

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

³ Department of Ecology & Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA

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

Current Opinion in Behavioral Sciences 2018, 21:27–32

This review comes from a themed issue on **The evolution of language**

Edited by **Christopher Petkov** and **William Marslen-Wilson**

<https://doi.org/10.1016/j.cobeha.2017.11.015>

2352-1546/© 2017 Elsevier Ltd. All rights reserved

Introduction

In human infants, much attention has been focused on the babbling period, where spontaneous streams of well-formed consonant-vowel syllables are the scaffold for simple words. While important, it is often overlooked that babbling is itself the culmination of the complex processes that make up prelinguistic vocal development. Early vocalizations, like cries, laughter, fussing, and cooing, are the infrastructure for babbling [1]. From a purely acoustic perspective, the increase in complexity from early to later vocalizations is continuous [2], and feedback

from caregivers is an instrumental driving force that can influence the maturation rate of prelinguistic vocalizations [3]. For example, the volubility of infants is influenced by social context [4], and caregivers who contingently respond to infant vocalizations spur the development of more complex vocalizations from those infants [5,6**]. Importantly, during prelinguistic vocal development, there is also growth of the vocal apparatus (the larynx, the vocal tract, and lungs) [7–9] (Figure 1a).

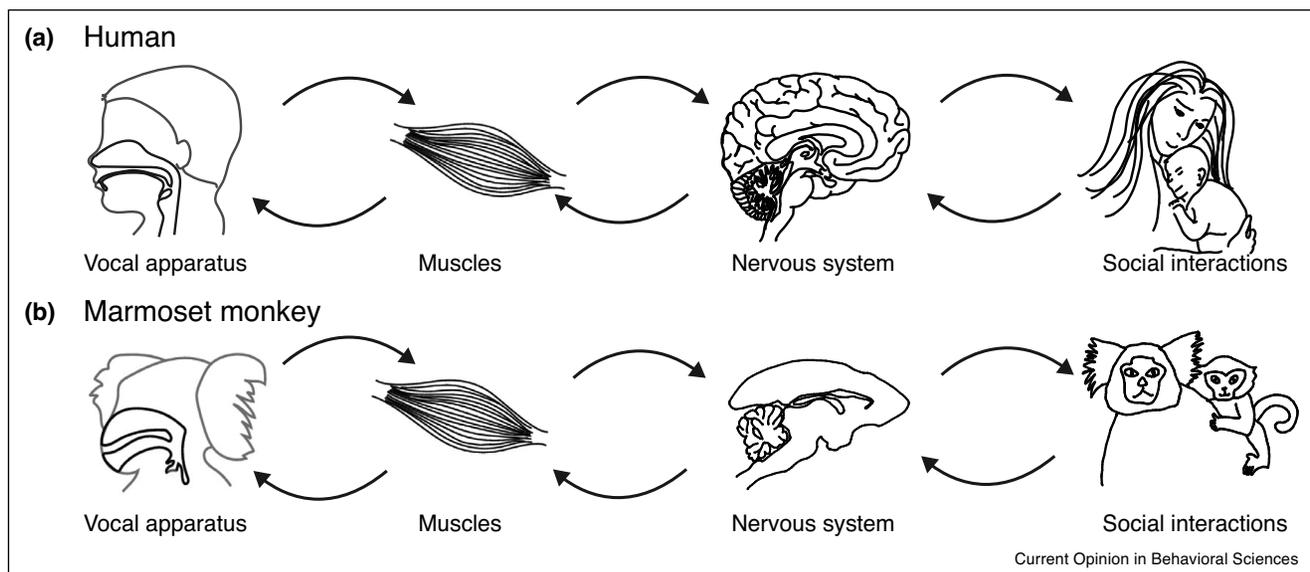
Models of vocal development that focus solely on the neural bases of learning often ignore these latter constraints. In order to understand the mechanisms underlying these parallel and interactive vocal developmental processes of growth and experience, we need an animal model system that shares these features with human development.

The marmoset monkey model system

Marmoset monkeys are a New World primate species and are cooperative breeders. Both parents, as well as older siblings and non-kin, will help care for offspring. This type of behavior is very rare among primates: only humans and members of the taxonomic group that includes marmosets (the Callitrichid family) exhibit this cooperative reproductive strategy. Thus, in terms of comparative developmental studies among human and nonhuman primates, marmoset monkeys (and other members of the callitrichid family) are a more compelling model system than the phylogenetically closer, but socially dissimilar, Old World apes and monkeys [10]. These cooperative breeding behaviors by humans and marmosets pave the way for more general prosocial cognitive processes [11,12], including those related to vocal communication [13].

Especially when compared to Old World primates like macaques, marmoset monkeys are quite agile in their vocal output. They and other closely related species readily adjust the timing of their contact calls to the timing of conspecific calls [14–16]; they also cooperatively modify the amplitude of their calls during vocal exchanges in accord with the perceived distance from conspecifics [17]. 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 (a behavior distinct from simple call-and-response behaviors observed among mates or competitors in other vertebrate species) [16]. Thus, while other nonhuman primates may

Figure 1



Vocal development in both humans **(a)** and marmoset monkeys **(b)** is an on-going, interactive process between the developing body, nervous system and social experience. One cannot ‘explain’ any aspect of vocal development without accounting for this interplay.

exhibit call-and-response behaviors with mates or specific group members (e.g., gibbons [18]; squirrel monkeys [19]; capuchins: [20]), this is not the same as turn-taking. Turn-taking is an on-going interaction with any conspecific involving multiple bouts of call-and-response exchanges. The turn-taking behavior by marmosets has the same coupled oscillator properties as human conversational turn-taking (but on a different time scale) [21,22]. In humans, this turn-taking behavior serves as a learning mechanism during prelinguistic vocal development: parents provide contingent responses to their offspring to spur the development of an infant’s vocalizations [5,6^{**}]. As we will describe below, marmoset monkeys have also adopted this social reinforcement strategy during their vocal development.

Vocal development in marmoset monkeys

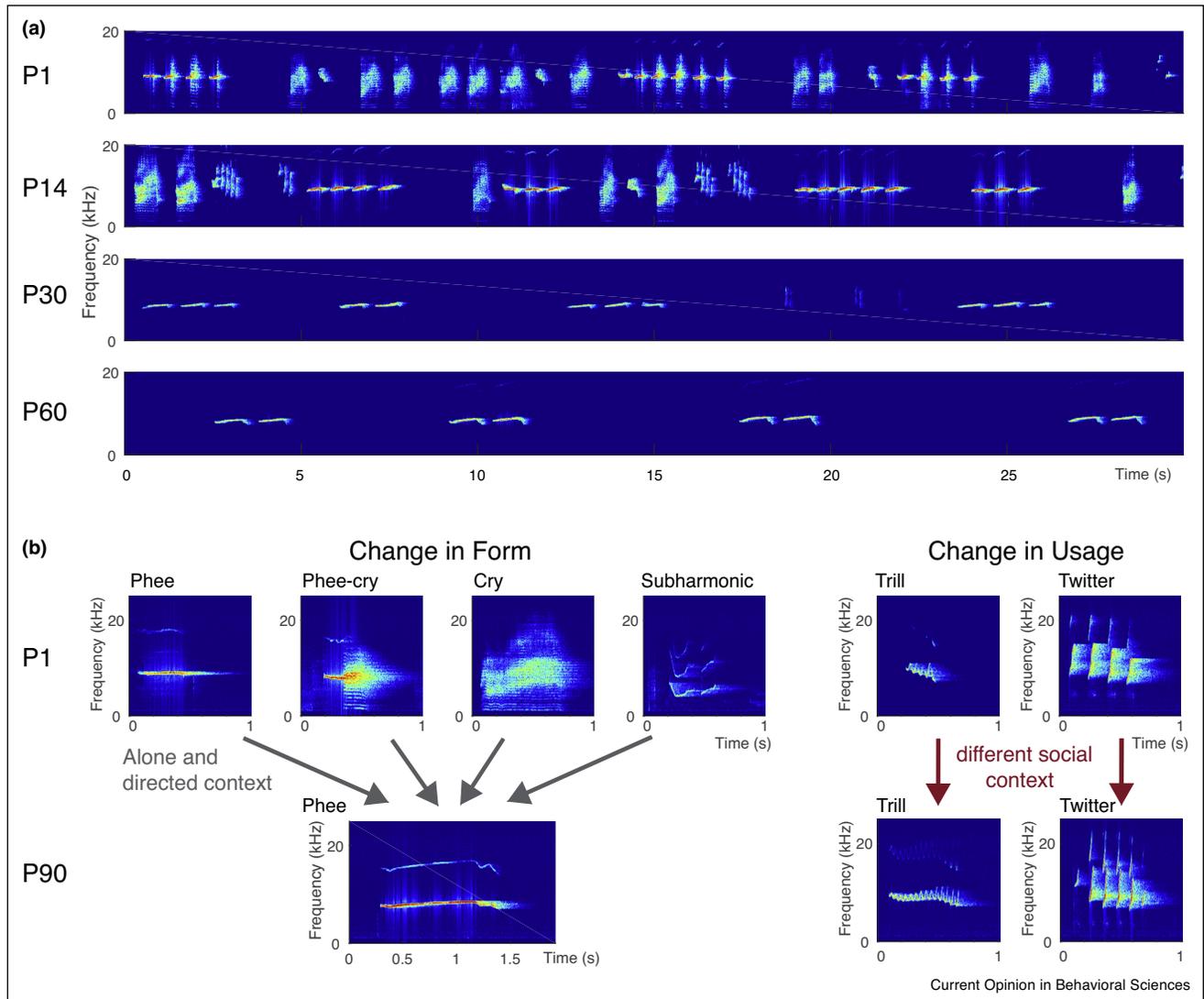
Contrary to what is known so far for other nonhuman primates [23], marmoset infants exhibit vocal learning in the first few months of postnatal life [24–26]. Similar to human infants [1,27] and juvenile songbirds [28^{*}], infant marmosets produce bouts of mature (‘twitters’, ‘trills’, and the contact call ‘phee’) and immature (‘cries’, ‘subharmonic phee’, ‘phee-cries’) vocalizations [24,25]. By two months of age, however, they only produce the appropriate contact calls (‘phee’) in both the undirected (alone) and directed contexts (vocally interacting with an out-of-sight conspecific) (Figure 2a). This suggests that two different vocal learning processes are at work: change in usage [26,29] and transformation of immature calls into mature versions [25] (Figure 2b). Twitters and trills are

frequently produced by marmosets of all ages [30,31], 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. It is worth noting, however, that this process has not been systematically and quantitatively examined in marmosets of any age and thus no specific pattern of vocal ontogeny for these calls have been identified.

Conversely, cries, phee-cries and subharmonic-phees are only produced by infants and are immature versions of the contact call [24,25]. Like the vocal transformations observed in preverbal human infants [2,32] and songbirds [28^{*}], these immature calls eventually become mature-sounding contact calls [24,25]. Measurements of several acoustic features [28^{*}] — duration, dominant frequency, amplitude modulation (AM) frequency, and Wiener entropy (a measure of noisiness) — revealed that marmoset monkeys go through this vocal transformation of their contact calls within their first two months of postnatal life [25]. During this time, contact calls lengthen in duration, decrease in dominant and AM frequencies, and decrease in entropy (i.e., they get more tonal).

Infants, in general, experience massive changes to their body morphology. Human infants, for example, double in weight within the first six months of their life (Center For Disease Control Data, USA); infant marmosets double their weight in about one month [25]. As the infants grow generally, so too do their vocal folds. Bigger vocal folds tend to oscillate slower leading to the

Figure 2



Infant marmoset vocalizations undergo dramatic acoustic changes. **(a)** Spectrograms of the vocalizations from one infant over two months of postnatal development. **(b)** Cries, phee-cries, and subharmonic-pees transform into mature contact calls whereas trills and twitters undergo a change in usage.

production of lower frequency vocalizations. Concurrently, the vocal tract (the mouth and nasal cavities) is lengthening, changing its resonance properties and affecting which frequency bands are amplified relative others in vocalizations [33,34^{*}]. Can these growth related changes account for the developmental trajectory of marmoset contact calls (as measured by the four acoustic parameters)? A comparison of the body weight curve of marmoset infants with the trajectory of acoustic change revealed that growth could explain a portion of the change in each acoustic parameter but none in its entirety [25]. Thus, growth constraints are important to marmoset vocal development, but are not the complete story.

Turn-taking as the developmental system upon which infant marmoset vocalizations are learned

The timing of when immature contact calls transform into mature-sounding versions is also influenced by how often parents provide contingent vocal feedback to those infant calls [25,35^{**}]. 'Contingent feedback' consists of those calls produced by marmoset parents that follow their infant's vocalization by a few seconds (similar to the contingent turn-taking pattern used by two adult marmosets during vocal exchanges). Parent-infant vocal interactions in marmosets recorded and quantified in the directed context (where infants and their mother or father were in auditory, but not visual, contact) revealed that the

timing of the transition varied substantially across infants (~10 to 40 days) [25]. A parental influence to account for this variability could be via the number of adult vocalizations the infant has heard (an exposure account) or via the number of contingent responses from parents. The latter turned out to be the case: the number of contingent vocal responses from parents correlated with the timing of the infants' phee-cry transition [25].

These data suggest that developing marmoset monkeys — unlike every other nonhuman primate investigated thus far — may be vocal learners [36]. However, a viable alternative hypothesis is that marmoset parents are simply responding more to healthier infants who develop their vocalizations more quickly than others. To address this, an experiment was performed to explicitly test whether or not the amount of contingent vocal feedback can accelerate the rate at which marmoset infants begin producing mature-sounding contact calls [35**]. Since marmoset monkeys typically give birth to dizygotic twins [37], the influence of genetics and the perinatal environment on vocal development could be controlled for. Starting from the first postnatal day to two months of age, randomly assigned infants were provided different levels of contingent feedback using closed-loop, computer-driven playbacks of parental contact calls. Twins who received high levels of contingent feedback learned to produce mature-sounding contact calls faster than their twin receiving low feedback [35**]. These results unequivocally demonstrate that infant marmoset monkeys use social experience to learn how to produce their vocalizations.

While these data demonstrate that marmoset monkeys learn how to produce their contact calls via contingent feedback from parents, the study did not address whether there are any long-term consequences to more or less parental contact. To put it another way, it seems that while parental feedback could influence the rate of vocal development, all infants would eventually be able to produce normal vocal output (even if they had no feedback at all). This is the case for babbling in human infants — even deaf infants babble, but do so with a substantial developmental delay [38]. In marmoset monkeys, however, an investigation of the vocal output of two sets of offspring from the same parents — one set normally-reared, the other separated from parents — revealed that parental contact of some form is necessary for normal vocal development in marmosets [39**]. In contrast to normally-reared monkeys, marmosets with limited parental contact, and who were now over a year old (the developmental equivalent of a 12-year old human), still produced infant-like specific vocal behaviors [39**].

Vocal production is demanding, eliciting high metabolic costs [40]. In marmoset monkeys, mature contact call production is particularly energetically-demanding, as it

requires high tension of the vocal folds and strong and sustained respiratory power to produce long, multi-syllabic, loud, and tonal vocalizations [34*]. The data show that, as in humans, changes in bodily growth shape the acoustic change in developing vocalizations, and these acoustic changes, in turn, shape the communicative experience marmoset infants have with their parents (Figure 1b).

How does cooperative breeding relate to vocal learning?

In the evolution of human communication, a key transition occurred when humans began to interact cooperatively [41]. Care of infants is probably the most important context in which cooperation with unrelated individuals occurs, and there is a strong correlation between the reproductive success of mothers and the amount of infant care provided by others [42]. When caregiver attention is a limited resource, and when non-maternal caregivers may have higher threshold than mothers to provide care, evolution may select for vocal behaviors that help infants attract attention [43]. Babbling and other infant vocalizations attract caregivers and trigger contingent responses from them. It is been suggested that infant babbling evolved to exploit auditory predispositions in adult receivers [44]. The fact that parents of both human and marmoset infants are more likely to give contingent responses to infant vocalizations when those vocalizations sound more adult-like is consistent with this 'receiver predisposition' idea [45,46]. This suggests that a vocal learning mechanism may have evolved to speed up the production of mature-sounding vocalizations (i.e., those that exploit the receiver predispositions) using social feedback because such vocalizations are more likely to elicit caregiver attention.

An integrated account of flexibility and constraints during vocal development

Given that vocal development is a systems phenomenon, its understanding requires consideration at many levels: from the vocal apparatus and its many associated muscles to the nervous system and the differential sensory feedback from environmental and social interactions (Figure 1a,b). Each element modifies both itself and others over time [47]. To capture such adaptive coordination, theoretical methods can be used to place these phenomena in a quantitative framework. Data from developing marmoset monkeys combined with optimal control theory were used to generate a developmental landscape based on the Waddington metaphor [48]. The model progressively adds one factor (such as the biomechanics of the vocal apparatus) and infers what aspects of the behavior can or cannot be explained before adding another factor (e.g., the strengthening of the musculature). This approach underscores the fact that the nervous system and its interplay with experience does not function in isolation; they must typically process sensory data

and communicate with the developing body to generate appropriate behaviors. The resulting coupling with changing morphology and other physiological systems both constrains and enables the adaptive behaviors that such neural circuits can produce [49,50].

Conflict of interest statement

Nothing declared.

Acknowledgements

The National Institutes of Health (NINDS) R01NS054898 (AAG), a James S. McDonnell Foundation Scholar Award (AAG) and an NSF Graduate Fellowship (DAL) supported this work.

References

- Oller DK: *The Emergence of the Speech Capacity*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.; 2000.
 - Kent RD, Murray AD: **Acoustic features of infant vocalic utterances at 3, 6, and 9 months**. *J Acoust Soc Am* 1982, **72**:353-365.
 - 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.
 - 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.
 - 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.
 - Goldstein MH, Schwade JA: **Social feedback to infants' babbling facilitates rapid phonological learning**. *Psychol Sci* 2008, **19**:515-523.
- An excellent experimental study demonstrating how contingent caregiver feedback can systematically influence vocal production in human infants.
- Vorperian HK, Kent RD, Lindstrom MJ, Kalina CM, Gentry LR, Yandell BS: **Development of vocal tract length during early childhood: a magnetic resonance imaging study**. *J Acoust Soc Am* 2005, **117**:338-350.
 - Boliek CA, Hixon TJ, Watson PJ, Morgan WJ: **Vocalization and breathing during the first year of life**. *J Voice* 1996, **10**:1-22.
 - Titze IR: **Nonlinear source-filter coupling in phonation: theory**. *J Acoust Soc Am* 2008, **123**:2733-2749.
 - 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.
 - 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.
 - Snowdon CT, Cronin KA: **Cooperative breeders do cooperate**. *Behav Process* 2007, **76**:138-141.
 - Borjon JI, Ghazanfar AA: **Convergent evolution of vocal cooperation without convergent evolution of brain size**. *Brain Behav Evol* 2014, **84**:93-102.
 - 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.
 - 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.
 - Takahashi DY, Narayanan DZ, Ghazanfar AA: **Coupled oscillator dynamics of vocal turn-taking in monkeys**. *Curr Biol* 2013, **23**:2162-2168.
 - Choi JY, Takahashi DY, Ghazanfar AA: **Cooperative vocal control in marmoset monkeys via vocal feedback**. *J Neurophysiol* 2015, **114**:274-283.
 - Mitani JC: **Gibbon song duets and intergroup spacing**. *Behaviour* 1985:59-96.
 - Masataka N, Biben M: **Temporal rules regulating affiliative vocal exchanges of squirrel monkeys**. *Behaviour* 1987, **101**:311-319.
 - Digweed SM, Fedigan LM, Rendall D: **Who cares who calls? Selective responses to the lost calls of socially dominant group members in the white-faced capuchin (*Cebus Capucinus*)**. *Am J Primatol* 2007, **69**:829-835.
 - Levinson SC: **Turn-taking in human communication—origins and implications for language processing**. *Trends Cogn Sci* 2016, **20**:6-14.
 - Ghazanfar AA, Takahashi DY: **The evolution of speech: vision, rhythm, cooperation**. *Trends Cogn Sci* 2014, **18**:543-553.
 - Egnor SER, Hauser MD: **A paradox in the evolution of primate vocal learning**. *Trends Neurosci* 2004, **27**:649-654.
 - Zhang YS, Ghazanfar AA: **Perinatally influenced autonomic nervous system fluctuations drive infant vocal sequences**. *Curr Biol* 2016, **26**:1249-1260.
 - 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.
 - 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.
 - Locke JL: *The Child's Path to Spoken Language*. Cambridge, MA: Harvard University Press; 1995.
 - 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.
- This was the first study to quantitatively and systematically document the trajectory of vocal development in a songbird. It represents the gold standard of vocal learning studies.
- Seyfarth RM, Cheney DL: **Vocal development in vervet monkeys**. *Anim Behav* 1986, **34**:1640-1658.
 - Bezerra BM, Souto A: **Structure and usage of the vocal repertoire of *Callithrix jacchus***. *Int J Primatol* 2008, **29**:671-701.
 - 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.
 - Scheiner E, Hammerschmidt K, Jurgens U, Zwirner P: **Acoustic analyses of developmental changes and emotional expression in the preverbal vocalizations of infants**. *J Phon* 2002, **16**:509-529.
 - Fitch WT, Giedd J: **Morphology and development of the human vocal tract: a study using magnetic resonance imaging**. *J Acoust Soc Am* 1999, **106**:1511-1522.
 - Teramoto Y, Takahashi D, Holmes P, Ghazanfar AA: **Vocal development in a Waddington landscape**. *eLife* 2017, **6**:e20782.
- This study provides a quantitative, systems biology framework for understanding and interpreting vocal development using the Waddington metaphor.
- Takahashi DY, Liao DA, Ghazanfar AA: **Vocal learning via social reinforcement by infant marmoset monkeys**. *Curr Biol* 2017, **27**:1844-1852.

This experimental study unequivocally demonstrates vocal production learning in infant nonhuman primates. It controls for genetics by using dizygotic twins and also shows that different acoustic parameters are more or less susceptible to contingent feedback from parents. This has the potential to better illuminate the underlying learning mechanisms.

36. Margoliash D, Tchernichovski O: **Marmoset kids actually listen.** *Science* 2015, **349**:688-689.
 37. Harris RA, Tardif SD, Vinar T, Wildman DE, Rutherford JN, Rogers J, Worley KC, Aagaard KM: **Evolutionary genetics and implications of small size and twinning in callitrichine primates.** *Proc Natl Acad Sci U S A* 2014, **111**:1467-1472.
 38. Oller DK, Eilers RE: **The role of audition in infant babbling.** *Child Dev* 1988:441-449.
 39. Gultekin YB, Hage SR: **Limiting parental feedback disrupts vocal development in marmoset monkeys.** *Nat Commun* 2017:8.
- This study shows that differential amounts of parental care have long term consequences on vocal output in marmoset monkeys. Thus, rather than vocal development ultimately achieving the same endpoint (but at different rates depending on the amount of parental interactions), this study shows that there are long lasting effects on vocal output as a function of parental care.
40. Gillooly JF, Ophir AG: **The energetic basis of acoustic communication.** *Proc R Soc Lond B: Biol Sci* 2010, **277**:1325-1331.
 41. Tomasello M: *Origins of Human Communication.* Cambridge, MA: MIT Press; 2008.
 42. Ross C, MacLarnon A: **The evolution of non-maternal care in anthropoid primates: a test of the hypotheses.** *Folia Primatol* 2000, **71**:93-113.
 43. Zuberbühler K: **Cooperative breeding and the evolution of vocal flexibility.** In *The Oxford Handbook of Language Evolution.* Edited by Tallerman M, Gibson KR. Oxford University Press; 2012:71-81.
 44. Locke JL: **Parental selection of vocal behavior: crying, cooing, babbling and the evolution of language.** *Hum Nat* 2006, **17**:155-168.
 45. 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.
 46. Takahashi DY, Fenley AR, Ghazanfar AA: **Early development of turn-taking with parents shapes vocal acoustics in infant marmosets.** *Philos Trans R Soc B* 2016, **371** 20150370.
 47. Byrge L, Sporns O, Smith LB: **Developmental process emerges from extended brain-body-behavior networks.** *Trends Cogn Sci* 2014, **18**:395-403.
 48. Waddington CH: *The Strategy of the Genes* . Routledge; 1957.
 49. Chiel HJ, Beer RD: **The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment.** *Trends Neurosci* 1997, **20**:553-557.
 50. Tytell ED, Holmes P, Cohen AH: **Spikes alone do not behavior make: why neuroscience needs biomechanics.** *Curr Opin Neurobiol* 2011, **21**:816-822.