From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Psychology

2017

AUTONOMIC NERVOUS SYSTEM DYNAMICS AND PRIMATE VOCAL PRODUCTI

Borjon, Jeremy Isaac

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01b5644v14s

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

AUTONOMIC NERVOUS SYSTEM DYNAMICS AND PRIMATE VOCAL PRODUCTION

JEREMY ISAAC BORJON

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF PSYCHOLOGY

ADVISER: ASIF A. GHAZANFAR

JUNE 2017

© Copyright by Jeremy Isaac Borjon, 2017.
All Rights Reserved

This dissertation, the publications it produces, and all my future endeavors are dedicated to:

My mother, Diana Borjon

My grandmother, Amparo Estrada Borjon

My tía, Esperanza Estrada

My tío, Raymundo Estrada (Sep. 7, 1924 - Jan. 4, 2014)

We did it.

Abstract

In humans, vocal production is typically cast as a cognitively intensive process, requiring a big brain and complex cognitive mechanisms. Recent studies have demonstrated that the vocal behavior of marmoset monkeys, a diminutive New World primate, is strikingly human-like in its dynamics and development. Considering the evolutionary gap between species, any shared neural mechanisms should be quite old. We propose that the convergent evolution of vocal behavior between marmosets and humans is partly due to neural mechanisms scaffolding upon a pre-existing arousal-regulation mechanism. We directly tested our hypothesis by recording natural changes in the adult marmoset's arousal state as it spontaneously vocalized. We demonstrate predictable changes in the animal's arousal state across multiple indices: heart rate, respiration rate, and motor activity. Moreover, the timing of vocal production in adult marmosets is coupled with the phase of a 0.1 Hz autonomic nervous system rhythm, the Mayer wave. The results of this study suggest that arousal dynamics are critical in assembling the timing and prediction of spontaneous primate vocalizations.

We next sought to determine whether the dynamics of the autonomic nervous system can act as a scaffold for the development of behavior. We densely sampled indices of arousal, motor activity, and vocal behavior from seven infant marmosets from birth until two months of age. Consistent with the human literature, changes in motor activity are predictive of vocal development in marmoset monkeys. Further, the developmental trajectory of the autonomic nervous system is itself oscillatory and correlates to changes in both the infant's vocal development and its motor development. We used statistical methods to extract causality between these processes. Changes in the development of the autonomic nervous system drives both the development of mature motor behavior and mature vocal production. Taken together, this dissertation argues that the autonomic nervous system is critical for the development and expression of spontaneous behavior in primates.

Acknowledgements

Chapters 1 & 3

These chapters were the result of many discussions with Asif A. Ghazanfar and Daniel Y. Takahashi. Portions of Chapter 2 were previously published with Asif A. Ghazanfar in *Brain, Behavior, and Evolution*. Portions of Chapter 3 will be published with Asif A. Ghazanfar, Daniel Y. Takahashi, and Yisi S. Zhang in the 39th Volume of *The Minnesota Symposium on Child Psychology*.

Chapter 2

I collected the data for the study described in Chapter 2 with the assistance of Daniel Y. Takahashi, Diego C. Cordero, and Diana A. Liao. I analyzed the data in collaboration with Daniel Y. Takahashi. This study was published in *The Journal of Neurophysiology* and the writing of the paper was done in collaboration with Asif A. Ghazanfar and Daniel Y. Takahashi.

Chapter 4

I collected the data for the study described in Chapter 4 with the assistance of Daniel Y. Takahashi, Diego C. Cordero, and Diana A. Liao. I pre-processed this data with the assistance of Anthony J. Sermarini. I analyzed the data in collaboration with Daniel Y. Takahashi. The manuscript for this study is under preparation for publication and being written in collaboration with Asif A. Ghazanfar and Daniel Y. Takahashi. A portion of this dataset was analyzed by Daniel Y. Takahashi and published in *Science*. A second portion of this dataset was independently analyzed and published in *Current Biology* by Yisi S. Zhang and Asif A. Ghazanfar.

Financial Support

During my first two years of graduate school, between 2012 and 2014, this work was supported by Princeton University's President's Fellowship. During my last three years of graduate school, from 2014-2017, this work was supported by the NSF Graduate Research Fellowship Program, which was awarded in 2012.

Editing

This dissertation was extensively edited and critiqued by Lydia Hoffstaetter. I am eternally grateful for her advice, her comments, and her patience.

List of Publications and Abstracts

Publications

Borjon J.I., Takahashi D.Y., Sermarini A.J., Ghazanfar, A.A. (In Preparation). The arousal dynamics of marmoset vocal development.

Borjon J.I., Takahashi D.Y., Cordero D.C., Ghazanfar, A.A. (2016). Arousal dynamics drive vocal production in marmoset monkeys. *Journal of Neurophysiology*. 116, 753-764.

Takahashi D.Y., Fenley A.R., Teramoto Y., Narayanan D.Z., **Borjon J.I.**, Holmes P., Ghazanfar A.A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science. 349*(6249), 734-8.

Borjon J.I., & Ghazanfar A.A. (2014). Convergent evolution of vocal cooperation without convergent evolution of brain size. *Brain, Behavior and Evolution*, *84*(2), 93–102.

Book Chapters

Ghazanfar A.A., Takahashi D.Y., Zhang Y.S., **Borjon, J.I.** (2017). Marmoset monkey vocal communication: Common trajectories with humans and possible mechanisms. In J. Ellison & M. Sera (Eds.). *The Minnesota Symposium on Child Psychology*, Vol. 39. Wiley Publishing.

Abstracts

Borjon J.I., Takahashi D.Y., Cordero D.C., Ghazanfar, A.A. (2015, October). Cardiorespiratory dynamics of vocal production in marmoset monkeys. Society for Neuroscience Annual Meeting, Chicago, IL.

Borjon J.I., Takahashi D.Y. and Ghazanfar A.A. (2013, October) Vocal-respiratory coupling in marmoset monkeys. Society for Neuroscience Annual Meeting, San Diego, CA.

List of Figures

Chapter 1

Figure 1.1 The dynamics of cooperative vocal behavior in the marmoset mimic the human timescale 18

Figure 1.2 Both marmosets and humans exhibit cooperative vocal behavior despite having disparate brain sizes. 20

Figure 1.3 Marmoset vocal communication can be modeled as a coupled oscillator 21

Chapter 2

Figure 2.1 Measuring respiration and heart rate from electromyographic (EMG) signal 45

Figure 2.2 Relationship between heart rate, respiration rate, and the production of vocalizations 47

Figure 2.3 Vocalizations are preceded by an increase in arousal 49

Figure 2.4 General motor activity partially ac- counts for fluctuations in arousal during a vocalization 51

Figure 2.5 Vocalizations are phase-locked to the rhythm of the autonomic system 53

Figure 2.6 Vocalizations are produced when perturbations of the system take the cardiorespiratory dynamics away from an attractor state 56

Chapter 4

Figure 4.1 The development of locomotor activity and vocal behavior in marmoset infants. 79

Figure 4.2 Motor behavior is correlated to vocal behavior in marmoset infants 81

Figure 4.3 Marmoset infants exhibit dramatic changes in arousal during development and these changes are correlated with vocal and motor development 83

Figure 4.4 Arousal development drives vocal behavior which drives the development of motor behavior 86

Contents

Abstract 6

Acknowledgements 7

List of Publications and Presentations 8

List of Figures 9

1 **Vocal Behavior in Humans and Marmosets: Converging Evolution**

1.1 Introduction 14

1.2 Cooperative communication: Vocal behavior in humans and marmosets 15

1.3 Why do humans and marmoset monkeys exhibit vocal behavior? 23

1.4 Conclusion 29

2 **The Arousal Dynamics of Marmoset Vocal Production**

2.1 Abstract 31

2.2 Introduction 31

2.3 Materials and Methods 34

2.3.1 Electromyography 34

2.3.2 Recording vocalizations 35

2.3.3 Heart rate analyses 36

2.3.4 Respiratory rate analyses 37

2.3.5 Call timing and probability measures 38

2.3.6 Index of motor activity 39

2.3.7 Population medians, individual medians and smoothing 40

2.3.8 Bootstrapped permutation tests and confidence intervals 40

2.3.9 Linear prediction of heart rate using respiration rate and the motion index 40

2.3.10 Probability density and power spectrum density estimates 41

2.3.11 Phase calculations 42

2.3.12 State space plot and multivariate Gaussian distribution fitting 42

2.4 Results 43
2.4.1 Increases in arousal precede the production of vocalizations and influence its probability and timing 43
2.4.2 The baroreceptor pathway does not wholly predict the link between heart rate changes and vocal production 46
2.4.3 The link between heart rate changes and vocal production is not wholly predicted by general increases in activity levels 50
2.4.4 Periodic autonomic nervous system fluctuations and vocal production: the Mayer wave 52
2.4.5 A dynamical systems view of vocal production 54
2.5 Discussion 55

3 **Vocal Behavior in Humans and Marmosets: Shared Development**
3.1 Introduction 62
3.2 Physiological influences on vocal development 62
3.3 Social influences on vocal development 66
3.4 The development of arousal dynamics 67
3.5 Conclusion 72

4 **The Arousal Dynamics of Marmoset Vocal Development**
4.1 Abstract 73
4.2 Introduction 73
4.3 Materials and Methods 75
4.3.1 General experimental procedures 75
4.3.2 Motor activity identification 76
4.3.3 Statistical methods 77
4.4 Results 78
4.4.1 Marmoset infants undergo dramatic changes in vocal behavior and motor activity across development 78
4.4.2 Marmoset motor development is correlated with vocal development 80
4.4.3 Marmosets exhibit dramatic changes in their arousal state across development and these changes are correlated to both vocal and motor development 82

4.4.4 Arousal development drives the development of motor activity, which drives vocal development 84

4.5 Discussion 87

5 **General Discussion** **91**

6 **References** **98**

Chapter 1

Vocal Behavior in Humans and Marmosets: Converging Evolution

1.1 Introduction

Successful communication through vocal production has the potential to facilitate cooperation between individuals via the coordination of behavior. Prior work aiming to understand the evolutionary basis of communication has generally cast the problem in terms of information theory: successful communication aims to transmit a signal to convey packets of information (Flack, 2013). Alternate work has cast the communication problem in terms of achieving coordinated behavior via manipulating the affective state of the receiver (Rendall & Owren, 2009). For example, one function of infant cries is to coordinate caregiving via a series of shrill, loud utterances. Implicit in both theories is an assumption that human-like communication requires a large brain and complex cognitive machinery. Recent work, however, has demonstrated that the dynamics of vocal behavior in the marmoset monkey a small New World primate, closely mimics that of humans (Takahashi, Narayanan, & Ghazanfar, 2013) and can be considered an instance of convergent evolution (Borjon & Ghazanfar, 2014). In this chapter, we will leverage the case study of the common marmoset to propose a third hypothesis in which the dynamics of cooperative vocal communication do not require complex cortical mechanisms. We will argue that the natural, spontaneous vocal production of the common marmoset is oscillatory in nature and may primarily serve as an arousal regulating mechanism. When paired with another conspecific, this oscillatory behavior becomes coupled, resulting in the emergence of cooperative vocal communication. We will propose putative mechanisms that may be involved in this oscillatory mechanism of vocal production and will continue to test this hypothesis in Chapter 2.

1.2 Cooperative communication: Vocal behavior in humans and marmoset monkeys

Humans orchestrate behavior through spoken language, gestures, and more often than not, a combination of speech and gestures. The evolutionary origins for this uniquely human form of communication are mysterious for many reasons. Primarily, there are few ways to reconstruct what ancestral vocalizations sounded like and how they were used. Moreover, soft tissues such as the brain and parts of the vocal apparatus generally do not fossilize. Relatedly, the dynamic social behavior that may have driven the evolution of more sophisticated communication is also difficult to reconstruct from the archaeological record. That said, we can infer from indirect evidence that early hominids were quite social. For example, evidence of hominid fire use dates to at least one million years ago (Berna et al., 2012) and evidence for a group of humans using fire in the form of a hearth has recently been pegged to at least 300,000 years ago (Shahack-Gross et al., 2014) . Whether our hominid ancestors gestured or spoke to one another (or did both) is an open question. Yet necessary and foundational to any instance of successful cooperative communication--no matter the modality or sophistication--is the capacity to structure the timing of produced vocalizations.

One way to enhance signal quality during communication is to prevent interference through the timing of vocal production. By pausing after transmission, a sender allows signals from other individuals to transpire and be heard before another signal is emitted. The elimination of overlap via sequential vocal production increases the likelihood of the signal being heard accurately. As a consequence, an exchange of signals between two or more individuals is ordered. A successful instance of human-like cooperative vocal behavior, for example, would involve person 1 speaking while person 2 attends, followed by a response from person 2, be it a statement or an indication for person 1 to continue speaking. The imposition of such timing rules during communication is a key feature of human social development (Jaffe et al., 2001; Jasnow & Feldstein, 1986). Mothers will use

pauses to establish rhythmicity in vocal, gestural, and gaze-driven interactions with their infant (Jaffe et al., 2001). The establishment of this rhythmicity has been demonstrated to result in measurable gains in word learning (Yu & Smith, 2012). Thus, we can envision the imposition of structure as a scaffold upon which a developing child can build mature communication strategies. Of course, not all conversations in humans adhere to strict rules. Face-to-face conversations between familiar individuals tend to be rapid with a substantial amount of overlap (O'Conaill, Whittaker, & Wilbur, 2009). However, between unfamiliar individuals or in unfamiliar contexts, or when individuals are out of view of one another, conversational exchanges solely via speech signals become more regulated (O'Conaill et al., 2009; Sellen, 1995). Thus, in general and throughout human development, structured vocal exchanges are a foundational principle upon which more flexible communication can occur.

Given its central importance in everyday human social interactions, it is natural to ask how vocal exchanges, a form of cooperation, evolved. It has been argued that human cooperative vocal communication is unique and evolved in, essentially, three steps. This argument was put forth most cogently by (Tomasello, 2008) but see also (Hewes, 1992; Rizzolatti & Arbib, 1998) for similar scenarios. First, an ape-like ancestor used manual gestures to point and direct the attention of others. Second, later ancestors with prosocial tendencies used manual gestures to mediate shared intentionality. Finally, and mysteriously, a transition from primarily gestural to primarily vocal forms of cooperative communication came about, perhaps in order to express shared intentionality more efficiently. Typically, no primates other than humans are thought to exhibit cooperative vocal communication. The implication being from this assumption is that human-like vocal communication requires a big brain and complex cognitive mechanisms.

Conversely, we hypothesize that cooperative vocal behavior could have evolved through a voluble and prosocial ancestor without the prior scaffolding of manual gestures or big brains. To test this

hypothesis, we studied the vocal exchanges (Takahashi et al., 2013)of a small New World primate found in South America: the common marmoset (*Callithrix jacchus*). Marmosets are part of the Callatrichinae subfamily of the Cebidae family of New World primates. The common marmoset is ~20 cm tall, weighs an average of 400g, and exhibits no sexual dimorphism in body size (Figure 1.1 A). Marmosets display little evidence of shared intentionality nor do they produce manual gestures. Like humans, they are cooperative breeders and voluble. Marmosets are among the very few primate species that form pair bonds and exhibit bi-parental and allo-parental care of infants (Zahed, Prudom, Snowdon, & Ziegler, 2008). These cooperative care behaviors are thought to scaffold prosocial motivational and cognitive processes such as attentional biases toward monitoring others, the ability to coordinate actions, increased social tolerance, and increased responsiveness to others' signals (Burkart & van Schaik, 2010; Snowdon & Cronin, 2007). Besides humans, and perhaps to some extent in bonobos (Hare, Melis, Woods, Hastings, & Wrangham, 2007), this suite of prosocial behaviors is not typically seen in other primate species.

When out of visual contact, marmoset monkeys and other callitrichid primates will participate in vocal exchanges with conspecifics (Ghazanfar, Smith-Rohrberg, Pollen, & Hauser, 2002; Ghazanfar, Flombaum, Miller, & Hauser, 2001; Miller & Wang, 2006). In the laboratory and in the wild, marmosets typically use "phees" (Figure 1.1 B,C), a high-pitched vocalization that can be monosyllabic or multisyllabic, as their contact call (Bezerra & Souto, 2008). A phee call contains information about gender, identity, and social group (Miller & Thomas, 2012; Norcross & Newman, 1993). As for humans, marmoset conversations occur spontaneously with another conspecific regardless of pair-bonding status or relatedness. Marmoset vocal exchanges can last as long as 40 minutes and have strikingly similar dynamics in timing to humans (Stivers et al., 2009; Takahashi et al., 2013). First, there are rarely if ever overlapping calls (i.e., no interruptions and thus no

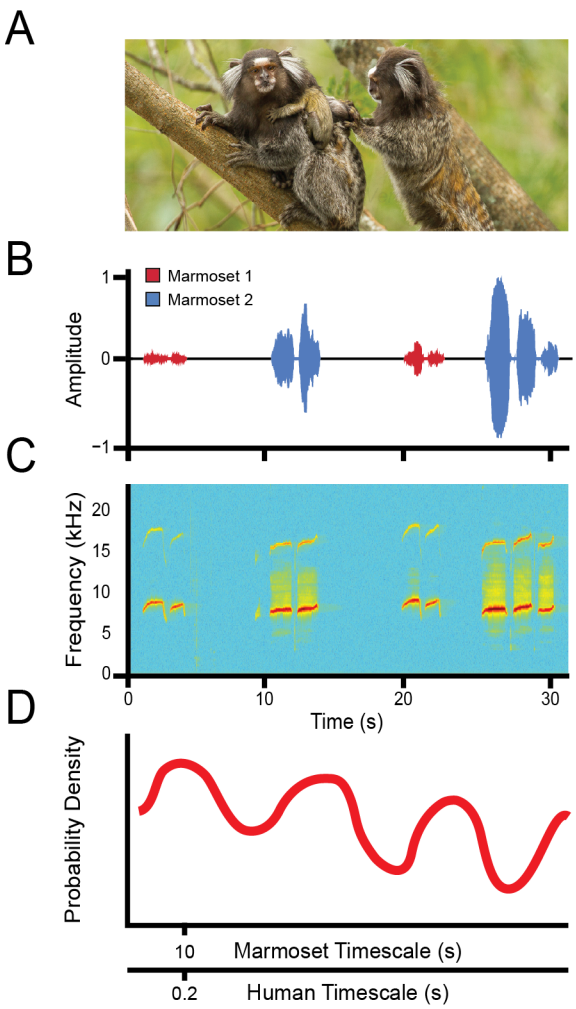


Figure 1.1 *The dynamics of cooperative vocal behavior in the marmoset mimic the human timescale.* A: A pair of common marmosets with two infants. (Copyright, Francesco Veronesi; used under Creative Commons license.) B and C: Waveform and spectrogram of an example phee call exchange. Used with permission from Takahashi et al. 2013. D: Schematic of a cross-correlation plot demonstrating the temporal structure underlying vocal exchanges in marmosets and humans. Note the difference in timescales, whereby the first peak of the cross-correlation is at 10 s for marmosets and 0.2 s for humans.

interference). Second, there is a consistent silent interval between utterances across two individuals (Figure 1.1 B,C). The similarities don't stop there.

In humans, dynamical system models incorporating coupled oscillator-like mechanisms are thought to account for the temporal structure of conversational vocal exchanges and other social interactions (Chapple, 1970; Oullier, de Guzman, Jantzen, Lagarde, & Kelso, 2008; R. Schmidt & Morr, 2010). In the vocal domain, such mechanisms have two basic features: 1) periodic coupling in the timing of utterances across two interacting individuals (Figure 1.1 D) and 2) entrainment, where if the timing of one individual's vocal output quickens or slows, the other's follows suit. The vocal exchanges of marmoset monkeys share both of these features (Takahashi et al., 2013). Thus, marmoset vocal communication, like human speech communication, can be modeled as loosely coupled oscillators. As a mechanistic description of cooperative vocal behavior, coupled oscillators are advantageous since they are consistent with the data from speech processing that brain oscillations are critical to temporal structure (Giraud & Poeppel, 2012) and its evolution (Ghazanfar & Takahashi, 2014b). Further, such oscillations do not necessarily require any higher-order cognitive capacities to operate (Borjon & Ghazanfar, 2014; Takahashi et al., 2013). In other words, a coupled oscillator can occur without the involvement of a big brain (Figure 1.2), something worth considering given the marmoset monkey's small encephalization quotient compared to great apes and humans (Jerison, 2012). A schematic of how two marmosets can be coupled to each other via vocalizations is illustrated in Figure 1.3. Consider first, a marmoset producing spontaneous vocalizations by itself with no conspecific responding (Figure 1.3 A). According to the model, preceding a vocalization, this marmoset experiences a drive (perhaps an increase in arousal due to social isolation) to produce a call. As the call is emitted, the sender hears its own signal and this auditory feedback inhibits the drive to continue vocalizing. Once the vocalization ends, the inhibition on the drive to vocalize is released and then, after some refractory period, the marmoset

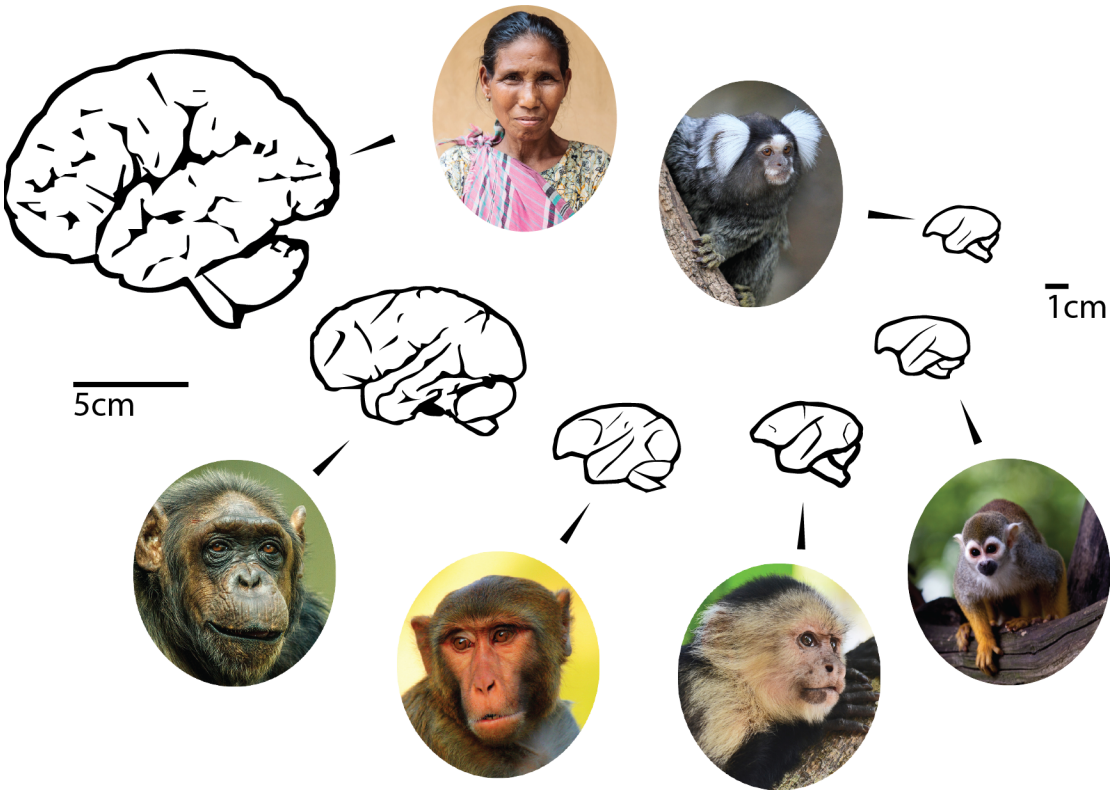


Figure 1.2 *Both marmosets and humans exhibit cooperative vocal behavior despite having disparate brain sizes.* A schematic demonstrating the variety of brain sizes in the primate order. To date, marmosets and humans are the only two primate species demonstrated to vocally cooperate via an extended series of vocalizations with any other conspecific. (Photographs used under Creative Commons license; photographers as follows: marmoset by Simon Harris; chimpanzee by S.B.J.; rhesus macaque by Stacey Osburn; capuchin by Dave Spangenburg; squirrel monkey by Tracy Lynn.)

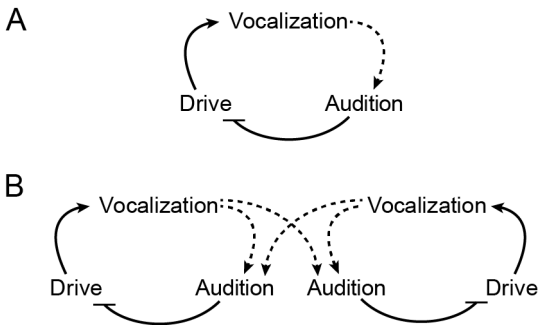


Figure 1.3 *Marmoset vocal communication can be modeled as a coupled oscillator.* A: An oscillator model of marmoset vocal communication while a marmoset is alone. B: By introducing a second marmoset, the vocalization of the first can alter the activity of the second, thereby creating a coupled oscillator.

produces another call. Thus, we have a simple oscillatory mechanism. Critically, the inhibitory influence of the auditory input on the drive to vocalize provides a natural mechanism for two marmosets to begin a vocal exchange and to vocally couple with each other and thereby cooperate at a systems level (Figure 1.3 B). A receiver marmoset out of sight from the sender hears the sender's phee call which inhibits its own drive to vocalize, preventing overlapping calls. Once the call ends, the drive to vocalize in the receiver is disinhibited and can hit threshold, whereby another call is initiated and the conversation continues. Now, we have a coupled oscillator. Since phee calls contain socially salient indexical information about the sender--and the structure of vocal exchanges prevent signal degradation caused by overlap--cooperation and social interactions are facilitated and reinforced. Overall, this model is consistent with the affective- and arousal-based theories of primate communication put forth by Owren and Rendall (Owren, Rendall, & Ryan, 2010; Rendall, Owren, & Ryan, 2009).

There is one major difference between human and marmoset monkey vocal exchanges: timescale. Across two individuals, the average human silent interval between utterances is ~240 ms, with a very large amount of variance (Stivers et al., 2009). In marmosets, this interval is much longer, ~3 - 5 s (Takahashi et al., 2013). This difference in timing can be explained in terms of "units of perception" (Ghazanfar et al., 2001). Since humans employ words as symbolic units of meaning, our minimal unit of perception in a conversation is on the order of a word or syllable (Chandrasekaran, Trubanova, Stillittano, Caplier, & Ghazanfar, 2009). Marmosets, on the other hand, may use the entire call, which may be multisyllabic and on the order of ~3-5 seconds (Takahashi et al., 2013). This is consistent with what is known about the units of perception in the contact calling behavior of the closely-related cotton-top tamarin (Ghazanfar et al., 2001). The proposed model in Figure 1.3 can work on either of these time-scales.

To be sure, many species of animals exchange vocalizations, but these usually take the form of a single "call-and-response" as opposed to an extended sequence of vocal interactions. For example, naked mole-rats (Yosida, Kobayasi, Ikebuchi, Ozaki, & Okanoya, 2007), squirrel monkeys (Masataka & Biben, 1987), Japanese macaques (Sugiura, 1998), large-billed crows (Kondo, Izawa, Watanabe, 2010), bottlenose dolphins (Nakahara & Miyazaki, 2011), and some anurans (Grafe, 1996) are all capable of simple call-and-response behaviors. However, there are some instances of extended, coordinated vocal exchanges as well. The chorusing behaviors of anurans and insects are the result of competitive coordination between neighboring males in an attempt to attract females (Greenfield, 1994). Another form of vocal coordination, duetting, occurs between pair-bonded songbirds (Logue, Chalmers, & Gowland, 2008) and gibbons (Mitani, 1985). Like the marmoset monkey vocal exchanges, these too indicate that a high-level cognitive capacity is not necessary for cooperative vocal behavior. However unlike marmosets and humans, these types of vocal coordination occur within the limited contexts of competitive interactions or pair-bonds. This inability to flexibly exchange vocalizations across conspecifics, regardless of pair-bonding status or relatedness, stands in stark contrast to the cooperative vocal behavior in marmosets and humans, who frequently initiate and sustain vocal exchanges with any conspecific (Takahashi et al., 2013). The frequency of interacting with non-related and non-pair-bonded individuals may be key to understanding the neural circuitry underlying both vocal production and cooperative vocal behavior.

1.3 Why do humans and marmoset monkeys exhibit cooperative vocal behavior?

Despite the stark contrasts in brain size, marmoset monkeys and humans can vocally cooperate with any conspecific regardless of pair-bonding status or relatedness. In other primates, vocal cooperation is restricted to certain types of conspecifics and is rarely cooperative. As some 40 million years have passed since the Old World and New World primate lineage split (Steiper & Young, 2006), we argue

that vocal cooperation arose as an instance of convergent evolution of behavior but perhaps through the activation of a shared (homologous) neuronal network. To unpack this argument, we will first examine the mechanisms driving convergent evolution in general then examine which pressures are shared between marmosets and humans. We will then argue that neuromodulatory changes to a common neural circuit could serve as a viable explanation for the convergent evolution of vocal cooperative behavior in marmosets and humans.

Natural selection acts on the behavior of an individual, not on isolated neural structures (Padberg et al., 2007). Evolutionary changes to organisms occur as parallel processes, involving the interplay between the environment, an individual's musculoskeletal structures, and the brain (Krubitzer & Kaas, 2005). In fact, given similar pressures, identical cortical areas can emerge by virtue of the constrained nature of cortical development (Krubitzer & Kaas, 2005; Padberg et al., 2007). For instance, cortical areas 2 and 5, associated with motor planning and coordination, are very well developed in macaques, an Old World monkey, as well as in cebus monkeys, a New World monkey. In other New World primates, however, areas 2 and 5 are either absent or poorly developed (Padberg, Disbrow, & Krubitzer, 2005). The reason for this becomes apparent when we consider that, of the New World primates, cebus monkeys are the only species known to use a precision grip—a grip that is common among Old World monkeys and apes, including humans (Padberg et al., 2007). Such a grip alters the mechanics of the hand, facilitating object manipulation through an increase in the possible number of digit configurations. In the context of homologous neuro-developmental trajectories, convergent evolution of biomechanics results in a constrained and predictable change in neural circuitry. Through this process, we see the emergence of identical cortical areas, in this case areas 2 and 5, across two species separated by a common ancestor 40 million years ago. It is possible that similar processes may be at play in marmosets and humans

whereby they exhibit homologous neuro-developmental trajectories that are influenced by convergent features of their socioecological environment—specifically, prosociality and volubility.

Cooperative breeding, a prosocial behavior, is only found in about 3% of mammals (Hrdy, 2005). Of those mammals, callitrichids are the only nonhuman primates known to exhibit this strategy (Burkart, Hrdy, & van Schaik, 2009; Hrdy, 2005). For marmosets, the rearing of infants is greatly reliant on a concerted effort among the breeding female, breeding male, non-breeding siblings and occasionally other familiar but unrelated group members (Burkart & van Schaik, 2010; Goldizen, 1990; Hrdy, 2005). Marmoset caregivers actively and frequently provision food for offspring (Burkart & van Schaik, 2010; Yamamoto & de Araújo Lopes, 2004). In contrast, infants of other New World monkeys, such as capuchins, rarely receive non-maternal care, and other group members seldom share food with them. Moreover, marmosets forage while carrying infants, and often compete to carry infants (Santos, French, & Otta, 1997; Snowdon & Cronin, 2007). The result of this foraging experience are gains in social learning by infant marmosets. Infant marmosets are highly unlikely to consume novel food items when away from adults unless the infants observed an adult eating the same food (Vitale & Queyras, 1997; Voelkl, 2006; Yamamoto & de Araújo Lopes, 2004). In contrast, capuchin infants do not exhibit any similar social learning preference: they are just as likely to consume novel foods when they are away or with adults (Fragaszy, Visalberghi, & Galloway, 1997). This cooperative breeding framework, in which non-parents within a social group spontaneously care for offspring other than their own has been argued to drive uniquely human cognition (Burkart et al., 2009). Experimentally, there is support for the notion that cooperative breeding leads to gains in social cognition (Burkart & van Schaik, 2010). When compared to their larger, independently breeding sister taxa Cebidae (which include squirrel monkeys and capuchin monkeys), marmosets exhibit greater socio-cognitive performance and much greater evidence of prosociality (Burkart & van Schaik, 2010).

Typically, benefits in socio-cognitive performance are thought to be the result of an increase in the size of the brain, specifically neocortex (Dunbar, 2009). This is not the case with the marmoset monkey. Marmoset brains are approximately 4 times smaller (Figure 1.2) than the closely-related New World squirrel monkeys and 6 times smaller than capuchin monkeys (Herculano-Houzel, Collins, Wong, & Kaas, 2007). However, in marmoset monkeys, selective pressures on body size may have led to a more efficient brain. Despite their small overall brain size relative to other primates, marmosets possess three times as many cortical areas as rats that exhibit a similar body size. To put it another way, a marmoset brain is five times larger than the brain of a rat; the brain of a marmoset is 2.7% of its body size, which is in similar proportion to humans (Stephan, Baron, & Schwerdtfeger, 2012), while the brain of the rat is less than 1% of its own body weight (Nieuwenhuys, Donkelaar, & Nicholson, 2014). Thus, these prosocial behaviors and, by extension, cooperative vocal communication must be mediated by the particular form and/or modulation of neural circuits that are not simply a consequence of increasing brain size.

Indeed, differential neuromodulatory influences on identical neural networks can result in very different behavioral outcomes (Katz, 2011). Take for instance maternal behavior in rats. In female rats, the medial preoptic area (MPOA) is critical for the appearance of maternal behavior especially when activitated by the hormone, estrogen (Numan, 1994). A lesion to the MPOA of a lactating female rat will cease any maternal behavior (P. Gray & Brooks, 1984). In the wild and in the laboratory, male rats do not readily exhibit maternal care to their pups; however, following the elimination of testosterone (via castration) and exposure to newborn pups, male rats begin to exhibit maternal care (Bridges, Zarrow, Goldman, & Denenberg, 1974). Importantly, this normally-latent male "maternal" behavior is also mediated by the MPOA, as lesions to the MPOA will result in the termination of the maternal behavior (Rosenblatt, Hazelwood, & Poole, 1996). Therefore, the only difference between male and female rats in regards to maternal care may be the presence of

testosterone at the MPOA. What this tells us is that, in the course of evolution, it is easier to modify existing neural pathways rather than create new ones. In light of this, we suggest that the variety of ways in which vocalizations are exchanged in primates - from the unrestricted nature of human and marmoset vocal cooperation to the restricted duetting of gibbons and the call and responses of squirrel monkeys - suggests that a shared neural connectivity homology may exist but is activated in different ways and/or to different degrees. Perhaps, this inconsistent utilization across closely related species implies variation in neuromodulatory action on the very same or similar circuit (Katz, 2011). Below, we conclude with a discussion of the possible neuromechanical underpinnings of vocal cooperation.

There is a large overlap in structures considered to underlie social behavior (and vocal behavior, in particular) as well as those underlying motivation and emotional arousal (Cardinal, Parkinson, Hall, & Everitt, 2002). The shared neural architecture enables a pathway in which social interaction is "rewarded" via the attachment of motivational value through arousal levels (Syal & Finlay, 2011). That is, fluctuating arousal levels could be a driving factor in vocal behavior. As cooperative vocal behavior is inherently rhythmic, it is not unreasonable to consider that this behavior exapted onto currently existing rhythmic mechanisms such as respiration, heart rate, or rhythmic fluctuations in neuroendocrine levels. Changes in these rhythmic mechanisms alter the autonomic nervous system, leading to changes in arousal (Shahid, Rahman, & Pilowsky, 2012). While many neuromodulators likely contribute to the production of a vocalization, here's one illustrative example. Orexin is a neuropeptide released by the hypothalamus and involved in the autonomic modulation of breathing and heart rate (Nattie & Li, 2012). Orexin knockout mice possess a lower basal blood pressure compared to wild types (Kayaba et al., 2003) and intracerebroventricular injection of orexin results in an increase in heart rate and blood pressure (Samson, Gosnell, Chang, Resch, & Murphy, 1999) as well as respiration (W. Zhang, Fukuda, &

Kuwaki, 2005). Intriguingly, neurotoxic lesions and chemical blocking of regions of the hypothalamus dense in orexin neurons results in a reduction of vocalizations as well as a lowering of blood pressure and pulse during contextual fear responses (Chen, Li, & Kirouac, 2014; Furlong & Carrive, 2007). As orexin is secreted in the brain, there are projections and receptors sensitive to this peptide throughout. In the rat, projections from the arcuate nucleus of the hypothalamus to anterior cingulate cortex (ACC) are known to exhibit orexigenic and anorexigenic phenotypes (Kampe, Tschöp, Hollis, & Oldfield, 2009). The ACC is one particular neural area in which vocal production and arousal intertwine. In cats, rats, and monkeys, electrical stimulation of ACC produces vocalizations (Devinsky, Morrell, & Vogt, 1995; Paus, 2001). These evoked vocalizations express an internal emotional state (Lewin & Whitty, 1960) and, thus, we can consider it one of the "drives" to vocalize (Figure 1.3 A). Importantly, the human ACC is involved in speech production (Sörös et al., 2006).

Is there evidence for a role of arousal on speech production in humans? Due to the presumption that speech is a purely neocortically-based phenomenon, it is typically argued that human communication is a behavior distanced from the state-governed vocalizations emitted by our non-human primate relatives (Owren et al., 2010; Rendall et al., 2009). This is a puzzling notion, especially if one takes into consideration the clinical psychology literature demonstrating a relationship between autonomic dysfunction and language production. For example, one characteristic symptom of anxiety disorders and mania are episodes of rapid speaking, often termed "pressured speech" (Pereira, Andreatini, Schwarting, & Brenes, 2014; Van Kammen & Murphy, 1975; Willson et al., 2005). Manic episodes usually entail a measurable change in autonomic sensitivity (Henry, Minassian, Paulus, Geyer, & Perry, 2010), suggesting that a dysfunctional autonomic system underlies mania (Latalova et al., 2010; Levy, 2013; McMeekin, 2002). Indeed, those with affective disorders often demonstrate an enlargement of the third ventricle, suggesting

dysfunction to the hypothalamic nuclei governing autonomic processes (Bhadoria et al., 2003). Separate from the clinical literature, speech motor coordination is influenced by fluctuations in the autonomic nervous system in both adults and children (Kleinow & Smith, 2006). Further, an increase in arousal is often accompanied by measureable changes in vocal prosody (Wiethoff et al., 2008). Changes in affect are indicated by changes to the fundamental frequency of the voice (McRoberts, Studdert-Kennedy, & Shankweiler, 1995). Decelerations in heart rate have also been demonstrated after speech production in typical humans (Peters & Hulstijn, 1984). Thus, human vocal communication, including speech, is greatly influenced by the autonomic system and fluctuating levels of arousal.

We propose that vocalization-induced reductions of arousal evolved into the mechanism underlying cooperative vocal behavior in marmosets and perhaps humans as well. In the marmoset monkey, auditory contact through phee calls has been shown to cause a reduction in cortisol levels (Rukstalis, Fite, & French, 2003). Unmodulated arousal is known to be damaging and even fatal to primates (Uno, Tarara, Else, & Suleman, 1989). Thus, modulating arousal levels and cooperative behaviors are intertwined (Figure 1.3): both producing and hearing vocalizations reduces arousal levels (Rukstalis et al., 2003). Consistent with this idea, it has been posited that human conversations play the same role as social grooming seen in apes and monkeys (Dunbar, 1998), suggesting that, in a sense, humans and marmosets have achieved "grooming at a distance" (Takahashi et al., 2013).

1.4 Conclusion

Here, we have examined an example of convergent evolution of vocal behavior between the common marmoset and human. We propose that the emergence of vocal cooperation between these distinct species is due to an evolutionarily old mechanism. We argue that one driver of vocal production is the regulation of arousal and that this regulation drives the observed oscillatory nature

of vocal production in the marmoset. When paired with a conspecific, this oscillatory vocal production serves as a foundation for cooperative vocal behavior to emerge. Such a mechanism stands in stark contrast to the traditional argument that cooperative vocal communication arose after being scaffolded onto manual gestures.

In the next chapter, we will directly test our hypothesis by investigating whether the natural, spontaneous vocal behavior of marmoset monkeys possesses any relation to fluctuations in arousal.

Chapter 2

The Arousal Dynamics of Marmoset Vocal Production

2.1 Abstract

Vocal production is the result of interacting cognitive and autonomic processes. Despite claims that changes in one interoceptive state—arousal—govern primate vocalizations, we know very little about how such states influence their likelihood and timing. Here we investigated the role of arousal during naturally occurring vocal production in marmoset monkeys. Throughout each session, naturally occurring contact calls are produced more quickly, and with greater probability, during higher levels of arousal, as measured by heart rate. On average, we observed a steady increase in heart rate 23 seconds before the production of a call. Following call production, there is a sharp and steep cardiac deceleration lasting approximately 8 seconds. The dynamics of cardiac fluctuations around a vocalization cannot be completely predicted by the animal's respiration or movement. Moreover, the timing of vocal production was tightly correlated to the phase of a 0.1 Hz autonomic nervous system rhythm known as the Mayer wave. Finally, a compilation of the state space of arousal dynamics during vocalization illustrated that perturbations to the resting state space increase the likelihood of a call occurring. Together, these data suggest that arousal dynamics are critical for spontaneous primate vocal production, not only as a robust predictor of the likelihood of vocal onset, but also as scaffolding, upon which behavior can unfold.

2.2 Introduction

Vocal production is the result of interacting cognitive and autonomic processes, similar to those underlying the production of facial expressions or other communicative signals (Gothard, 2014). In

nonhuman primates (hereafter, primates), we know that many external factors influence the timing and likelihood to produce a vocalization (Seyfarth & Cheney, 2003). Such factors include an individual's recent history of social interactions (Cheney & Seyfarth, 1997), strength of bonds (Kulahci, Rubenstein, & Ghazanfar, 2015), position in a social hierarchy (Kitchen, Cheney, & Seyfarth, 2005), distance from conspecifics (Choi, Takahashi, & Ghazanfar, 2015; Fischer, Hammerschmidt, Cheney, & Seyfarth, 2001), seeing a predator (Seyfarth, Cheney, & Marler, 1980), hearing conspecific calls in general (Ghazanfar et al., 2001; Takahashi et al., 2013), and hearing calls from a specific individual (Miller & Thomas, 2012). These findings are bolstered by recent neurophysiological studies implicating neocortical structures in the production of vocalizations (Coudé et al., 2011; Hage & Nieder, 2013; Hage, Gavrilov, & Nieder, 2013). While the above external factors have been exhaustively studied, we know very little about how internal, autonomic processes may influence vocal production in primates.

One function of the autonomic nervous system is to control the state of arousal. Arousal is relevant for a range of behaviors (Pfaff, 2006), and an animal would be said to exhibit a high arousal state if it is more alert to sensory stimuli, more motorically active and more reactive (Garey et al., 2003; Pfaff, 2006). In the domain of vocal behavior, the focus has been on linking presumptive arousal level changes to changes in behavioral context and, subsequently, to changes in the acoustic properties of vocalizations. For example, in tree shrews, decreasing physical distance between a male and female is thought to increase arousal levels in the female, resulting in a higher rate of vocal output with higher fundamental frequencies (Schehka, Esser, & Zimmermann, 2007). Along the same lines, levels of "response urgency" can be modulated by distance from a predator, and this is linked to changes in the acoustic structure of alarm calls in squirrels (Owings & Virginia, 1978), marmots (Blumstein & Armitage, 1997), and meerkats (Manser, Bell, & Fletcher, 2001). Acoustic changes in cat (Scheumann et al., 2012), bat (Bastian & Schmidt, 2008) and horse (Briefer et al.,

2015) vocalizations were also reported as a response to context-related arousal fluctuations. In primates, changes in arousal levels are also repeatedly linked to changes in vocal acoustics (Hammerschmidt & Fischer, 2008; Owren, Amoss, & Rendall, 2011; Rendall & Owren, 2009). In baboons, for instance, Rendall (Rendall, 2003) examined the grunt vocalizations of females in two different contexts and, within these two contexts, categorized low and high arousal states based on patterns of movement. He found that variation according to arousal levels was mainly associated with changes in the rate of calls produced and their duration, as well as changes in the fundamental frequency (Rendall, 2003).

In the current study, we address a different question with regard to arousal and vocal production: What mechanism drives an individual primate to produce a vocalization at all? To investigate this, we needed a model system that readily produces vocalizations without the necessity of an external event such as a conspecific vocalization, predator, or other overt sensory trigger. Such a context would allow us to focus solely on how changes in arousal may or may not be linked to vocal production. The calling behavior of marmoset monkeys (*Callithrix jacchus*) meets this criterion (Borjon & Ghazanfar, 2014). When alone, marmoset monkeys vocalize readily and spontaneously (Bezerra & Souto, 2008; Takahashi et al., 2013). Typically, if a conspecific is within hearing range, the two individuals will take turns exchanging calls (Takahashi et al., 2013), but if not, the isolated marmoset will continue calling. Moreover, in this undirected context, nearly all the calls they produce are a single type of contact call, the "phee call." It is a high-pitched vocalization with one or more largely identical "syllables" (Bezerra & Souto, 2008; Takahashi et al., 2013). As context controls the call type, this allows us the clearest possible means to investigate how arousal may influence vocal production.

We used electromyography to measure heart rate and respiration in adult marmosets in this undirected context. We had multiple hypotheses. First, we predicted that in this controlled context,

marmosets would experience fluctuations in their arousal levels (i.e., changes in heart rate) and that higher levels of arousal would lead to the production of vocalizations. Second, based on our previous behavioral study indicating that the spontaneous calling behavior of marmosets exhibits a periodicity of around 0.1 Hz (Takahashi et al., 2013), we predicted that arousal levels would also fluctuate at this same frequency and that vocal production would be phase-locked to this oscillation. Finally, we predicted that the relationship between our heart rate measure of arousal and vocal production was not 1) simply due to the tight link between heart rate and respiration (the driving force of vocal production) or 2) simply due to increased levels of movements just prior to vocal production.

2.3 Materials and Methods

The subjects used in this study were four adult common marmosets (*Callithrix jacchus*; three males and one female, average age was 5.1 yrs) housed at Princeton University. The median number of sessions they participated in was 7 sessions per animal (For monkeys 1-4, respectively: 9, 3, 9, and 5 sessions). All marmosets lived with their pair-bonded mates in family groups and were born in captivity. They were fed daily with standard commercial chow supplemented with fresh fruits and vegetables and had ad libitum access to water. The colony room was maintained at a temperature of approximately 27°C with 50%– 60% relative humidity, and a 12 hr light:12 hr dark cycle. All experimental sessions were conducted during daylight hours between 12:00 and 18:00 and were performed in compliance with the guidelines of the Princeton University Institutional Animal Care and Use Committee.

2.3.1 *Electromyography*

For the experiment, the animal was placed in a testing box. The testing box was made of Plexiglas and wire in a triangular prism shape (.30m x .30m x .35m). To record the EMG signal, we used two pairs of Ag-AgCl surface electrodes (Grass Technology). Tethered electrodes were sewn into a soft elastic band, which was clasped around the animal's thorax. One pair of electrodes was affixed to the chest area close to the heart and the second pair placed on the back, close to the diaphragm. To improve the signal-to-noise ratio, we applied ECl gel on the surface of each electrode. If needed, marmosets were shaved around the thorax. Each pair of electrodes was differentially amplified (250x) with the resulting signal sent to a Plexon Omniplex, where it was digitized at 40 kHz and sent to a PC for data acquisition. Recorded sessions were between 20 and 50 minutes (avg±sd for Monkeys 1-4, respectively: 21.86±11.70; 31.76±5.51; 22.63±12.32; 15.44±4.86).

2.3.2 *Recording vocalizations*

Vocalizations were recorded using a Sennheiser MKH416-P48 microphone suspended 0.9m above the testing box. The microphone signal was sent to a Mackie 402-VLZ3 line mixer whose output was then relayed to the Plexon Omniplex for acquisition. This ensured the timing accuracy of the vocal signals relative to the EMG signals. Although marmosets produce a number of distinct vocalizations in a number of different contexts (Bezerra & Souto, 2008) 99% of the vocalizations recorded in the undirected context of social isolation used here were phee calls, the species-typical contact call. We used the same criterion we established in previous work for computationally defining and segmenting phee calls and their syllables (Takahashi, Fenley, & Ghazanfar, 2016; Takahashi et al., 2015). This and all subsequent analyses were conducted in Matlab 2016b. The segmentation routine automatically detected the onset and offset of any acoustic signal that differed from the background noise at a specific frequency range. To detect the differences, we band-passed the entire recorded signal between 5 and 8 kHz for each session. This frequency band captures the

fundamental frequency of phee calls. We then compared the amplitude of the signal at this frequency band for the periods without calls and during a call. A simple threshold was enough to distinguish both periods. Onset-offset gaps longer than one second indicated separate calls, whereas gaps shorter than or equal to 750 ms indicated syllables from the same call. After this procedure, we manually verified for each call whether the automatic routine correctly identified multi-syllabic calls. Furthermore, we counted the number of syllables of each call, and if there was any mismatch between the automatic onset–offset detection and the call observed in a spectrogram, we marked the onset and offset of the call manually. A total of 427 calls were produced across a span of 26 sessions (For monkeys 1-4, respectively: 192, 29, 120, and 86 calls).

2.3.3 *Heart rate analyses*

Surface electrodes along both the ventral and dorsal thorax of the marmoset can each pick up both heart rate and respiratory signals. The strength of these signals can vary throughout a session as the positioning of the surface electrodes changes with the animal's movement. Therefore, we sought to screen the quality of the data and determine which channel exhibited the cardiac signal with the largest signal-to-noise ratio on a session-by-session basis. This channel was used for all heart rate analyses for that session. Because automated methods of detecting and removing noise from these selected channels were sub-optimal, we manually identified and isolated motion artifacts or signal cutoffs. To minimize bias, we did the following: For each session, signals from both EMG channels were divided into 10-second segments and signal pairs were presented in random order for visual inspection. Regions exhibiting signal loss or motion artifacts were replaced with NaNs (not-a-number, Matlab).

Following this screening, we down-sampled the data from 40 kHz to 1500 Hz to extract the cardiac signals. These signals were then high-pass filtered at 15 Hz to preserve the rapid waveform

of the heartbeat. To remove line noise, the resulting signal was notch filtered at 60 Hz. Heart beats were detected using an adaptive threshold of 1-s duration to find cardiac spikes greater than the 95th percentile of the amplitude at each second of the signal. Occasionally, the detection of heartbeats would pick up a spurious spike close to the actual heartbeat. This would result in a series of spikes occurring with an inter-spike timing faster than 100ms, corresponding to a biologically implausible 600 beats per minute. We therefore implemented a speed threshold, whereby two spikes occurring in succession quicker than 100 ms were substituted with a single spike located at the midpoint. It is also possible for the EMG recording to miss a heartbeat. If any inter-beat interval between spikes was larger than 400 ms (150 beats per minute), which is the minimum heart rate reported in conscious, unrestrained adult marmosets (Schnell & Wood, 1993), we replaced the intervening signal with NaNs to account for the signal loss. To calculate heart rate, we constructed a binary series of heartbeat counts and convolved the resulting series with a 1s width Gaussian window. This method is often used in estimating the instantaneous rate of action potential occurrences in neurophysiological data (Shimazaki & Shinomoto, 2010).

2.3.4 *Respiratory rate analyses*

The same noise control procedure used to determine the best channel of cardiac activity was used to determine which channel demonstrated the best respiration signal. Segments of signals that did not demonstrate a respiratory signal were excluded from further analysis by replacing them with NaNs. Since we were interested in physiological signals related to respiration (< 3 Hz), we downsampled the data from 40 kHz to 50 Hz. To remove the DC offset, we high-pass filtered the signal at 1 Hz. This resulting signal was further low-pass filtered at 4 Hz to isolate respiratory activity. Because mammals, including primates, produce vocalizations using expiration-driven vocal fold movements (Fitch & Hauser, 1995), we wanted to make sure that the respiratory signal always had a consistent

phase relationship with vocal output. That is, we wanted to keep the expiratory phase of the respiration signal in a consistent orientation relative to vocal production (so that the beginning of expiration was the peak of the respiratory curve). To do this, we ensured that the slope between the point of respiration at call onset and the point of respiration at the peak call amplitude was negative. The phase coordinates of the respiratory signal were then determined by calculating the angle of the Hilbert transform of the resulting signal. To calculate respiratory rate, we first detected the points in time where the angle of the Hilbert transform was 0, representing the beginning of expiration. These time points were converted into a binary representation and convolved with a 2 s Gaussian window. Again, this method is often used in estimating the instantaneous rate of spike occurrences (Shimazaki & Shinomoto, 2010).

2.3.5 *Call timing and probability measures*

We wanted to understand the relationship between heart rate and call timing on a session level. That is, we wanted to determine how quickly a call will occur given the current heart rate at any point during a session. To obtain the heart rate at a resolution of 1-second, we divided the data into 1-second bins, and the average heart rate for each bin was computed. There are considerable differences in heart rate between individuals (Cacioppo, Uchino, & Berntson, 1994). To normalize changes in the magnitude of heart rate between our marmoset subjects, we converted the 1-second averages into percentiles (from the 0th percentile up to the 100th percentile, in 1 percentile wide bins). Call timing relative to each bin of data was determined by measuring the interval from each bin to the onset of the closest call. To visualize the average probability curve, we fit the data to a cubic spline using csaps in Matlab.

We also wanted to know the probability of observing a call given a certain heart rate from any point within the session. Using the 1-second averages, we measured the probability of observing a

single call within 23 seconds of each bin. This time point corresponds to the minimum average time interval from each heart rate percentile to the closest subsequent call, as described in the paragraph above. To better visualize the probability of observing a single call within 23 seconds, we grouped the data at 5 percentile steps. To show the slope of the correlation, we used linear regression to plot the line-of-best-fit. We calculated the correlation using corr in Matlab and running a Spearman correlation between the averaged data points and heart rate percentiles.

2.3.6 *Index of motor activity*

To determine whether a marmoset was moving and active during the session, we analyzed the videos recorded during the experiment (23 out of 26 sessions were video recorded at 30fps using Plexon Cineplex). Each recording was split into segments of 30 frames. For each segment, we took the absolute difference between the 1st and 30th frame. This value corresponded to the difference in pixel luminance between the two frames. A higher value indicates more of a difference, signifying movement. We calculated, for each session, the 90th percentile of these pixel values and used that as a cutoff to construct a binary signal. Values higher than the cutoff were instances of movement, while values below the cutoff were not.

We manually scored the timing of each vocalization by watching the muted video. We only marked instances where we could clearly see, via mouth movements, the beginning and end of a vocalization. The beginning of a vocalization was marked at the frame prior to the animal opening its mouth. The end of the vocalization was marked at the frame after closure of the mouth. From this, we accounted for 239 vocalizations of the 427 used in heart rate and respiratory rate analysis in 23 videos. As the video frame rate is at 30 Hz, our identification of call types had a maximum resolution of 1/30 second. Using the video-marked vocalization onsets, we were able to compile data for Figure 2.4 B.

2.3.7 *Population medians, individual medians and smoothing*

To visualize the dynamics of heart rate, respiratory rate, and motor activity surrounding a call event, we took the median of 40-seconds before and after both call onset and offset, and we fit a cubic spline to the data using csaps in Matlab (smoothing parameter 0.3 for the population median and 0.01 for the individual medians).

2.3.8 *Bootstrapped permutation tests and confidence intervals*

To determine whether the changes in heart and respiratory rate and motor activity were significant, we performed a bootstrapped significance test. For each session, we chose random segments equivalent in number and length to the calls produced in that session from our dataset. The median of this randomly selected data was calculated and fit to a cubic spline using csaps in Matlab. We repeated this process 1,000 times. The 95 percent threshold for significance corresponds to the 2.5th and 97.5th percentiles of the bootstrapped medians.

Confidence intervals for the medians of heart and respiratory rate and motor activity were generated by randomly resampling, with replacement, from the signals used to calculate the plotted medians. We calculated the median of the resampled signals and fit it to a cubic spline using csaps in Matlab. We repeated this process 1,000 times. The 95 percent confidence interval corresponds to the 2.5th and 97.5th percentiles of the bootstrapped medians.

2.3.9 *Linear prediction of heart rate using respiration rate and the motor index*

We sought to linearly predict heart rate from ongoing respiratory activity. We first calculated the cross-correlation of heart and respiratory rate using the median of the dataset, and then took the maximal and minimal peak after 0 seconds delay. These points indicate the times at which

respiratory rate maximally correlates to heart rate. Aligned to call onset, the first positive peak occurs at 5 seconds while the negative peak occurs at 23.34 seconds. Aligned to call offset, the first positive peak occurs at 5.78 seconds while the negative peak occurs at 21.48 seconds. We built linear predictors of heart rate from respiratory rate starting from these points in time. To build them, we calculated the linear regression between heart rate and respiratory rate. We created an interval of predicted heart rate using a bootstrap procedure for the data aligned to call onset and call offset. For this, we uniformly and randomly sampled segments of heart rate and the corresponding respiratory rates. For each pair of segments, we then calculated the predicted heart rate curve using the associated respiratory rate curve. This resampling procedure was repeated 1,000 times. The 95 percent confidence interval corresponds to the 2.5th and 97.5th percentiles of the predicted heart rate.

We constructed an analogous linear predictor using the average of the motion index and the median of the heart rate data. As the motion index is gathered at a sampling rate of 1 Hz, we down-sampled the 50 Hz heart rate median to 1 Hz in order to match the motion index. We cross-correlated the down-sampled heart rate median to the mean of the motion index. Aligned to call onset, the first positive peak occurs at 6 seconds while the negative peak occurs at 29 seconds. Aligned to call offset, the first positive peak occurs at 8 seconds while the negative peak occurs at 32 seconds. We calculated the linear predictor sequence as described above. The 2.5th and 97.5th percentiles of the predicted heart rate were up-sampled to 50 Hz from 1 Hz. Down-sampling and up-sampling were conducted using the resample function in Matlab.

2.3.10 *Probability density and power spectrum density estimates*

To determine the oscillatory nature of marmoset vocal production, we first calculated the timing interval (in seconds) between spontaneous vocalizations for every marmoset. The interval was from

the onset of one call until the beginning of the subsequent call. Using this input calculated across all animals, we constructed a probability density function using ksdensity in Matlab.

To visualize the range of frequencies underlying the dynamics of the heart and respiratory rate signal, we calculated the power spectrum density of the signal 25-seconds prior to each call onset for every animal and session. Each signal was detrended and normalized by calculating its z-score. We used pburg in Matlab to calculate the power spectrum density estimates. To calculate the population power spectrum, we computed the average across all calls.

To calculate the call-rate, we constructed a binary equivalent in length to each of the 26 sessions. This binary was convolved with a 1-second bandwidth Gaussian window, then detrended and z-scored. We calculated the average power spectrum density for the resulting signals using pburg. To allow comparison between the signals, we normalized all power spectra to their first observed peak.

2.3.11 *Phase calculations*

To calculate the phase of the heart and respiration rate, we calculated the angle of the Hilbert transform for each call. We then calculated a circular phase density plot for the phase at call onset using circ_ksdensity in Matlab. A bootstrapped significance test was constructed by randomly sampling, for each session, 1,000 times, a number of points equal in number to the number of calls and calculating the 95th percentile of the phase location.

2.3.12 *State space plot and multivariate Gaussian distribution fitting*

To understand the cardiorespiratory dynamics underlying vocal production, we constructed the state space from the median heart rate and respiratory rate during a vocalization. We plotted heart rate against respiratory rate for 25 seconds before call onset until 20 seconds after the call begins. To

visualize the scope of cardiorespiratory variability within our dataset, we fit the permuted dataset used for the bootstrapped significance test of heart and respiratory rate variability to a Gaussian distribution. We first calculated the cardiorespiratory dataset's covariance matrix using cov in Matlab. We then fit a multivariate normal Gaussian distribution to the permuted data and calculated the contours encompassing the 95% confidence interval for the distribution.

2.4 Results

2.4.1 *Increases in arousal precede the production of vocalizations and influence its probability and timing.*

We used surface electromyography (EMG) to measure muscle activity around the thorax of four unrestrained adult common marmosets as they spontaneously produced contact "phee" calls in an undirected context (i.e., social isolation). This allowed us to capture ongoing heart rate simultaneously with respiratory activity, the latter of which is directly related to mammalian vocal production. The undirected context contained no external social signals, meaning there were no response phee calls produced by an out-of-sight conspecific, as would typically be the case in a contact-calling scenario (Takahashi et al., 2013). Figure 2.1 A shows the standard position of electrodes on the body and an example of the raw EMG signal recorded during a one-syllable phee call. An average power spectrum of the EMG signal across the entire dataset reveals two distinct peaks: a slow frequency associated with respiration and a higher frequency linked to cardiac activity (Figure 2.1 B). An adaptive threshold was used to identify cardiac spike trains in the high frequency band, and this was then used to measure heart rate changes. In the lower frequency band, respiratory peaks were identified using a Hilbert transform, and then converted into respiratory rate. As in humans (Cacioppo et al., 1994), marmosets exhibited considerable individual differences in the range of heart and respiration rates. To put these values in a uniform scale, we normalized each subject's

heart and respiratory rate into percentiles. Figure 2.2 A shows the different ranges of heart rate (avg±sd beats per minute (bpm): M1, 321±75; M2, 409±39; M3, 388±45; M4, 382±55) and respiratory rate (avg±sd respiratory cycles per minute (rpm): M1, 85±24; M2, 85±24; M3, 77±19; M4, 83±27).

If arousal contributes to spontaneous vocal production, we would predict an interaction between heart rate—a temporally precise index of arousal state (M. A. Gray et al., 2012; Obrist, Webb, Sutterer, & Howard, 1970)—and the timing and likelihood to vocalize. A total of 427 spontaneous phee calls were recorded during our EMG experiments. Figure 2.2 B shows the negative correlation between time of vocalization onset and heart rate for both the population and individual monkeys. The onset of call production was quicker following epochs of high versus low heart rates (for population, Spearman correlation, -0.978, $p<0.001$). The exact onset of call production relative to heart rate is also highly dependent on the individual animal, with some animals producing calls more frequently than others. To illustrate this, we have provided data for each animal in our study. While high heart rate modulates the timing of vocal onset, it does not provide information regarding the likelihood of a call occurring after a given level of cardiac activity. For example, it could be that at a heart percentile of 90, a vocalization is 50% likely to occur quickly, and at a heart rate percentile of 25, a vocalization is also 50% likely to occur but takes longer to be elicited. Figure 2.2 C shows a positive correlation between heart rate and the probability of the phee call being produced within 23 seconds of a given heart rate percentile (for population, Spearman correlation, 0.920, $p<0.001$). Thus, the higher the heart rate, the greater the probability of vocal output. These data show that both the timing and probability of vocal production are correlated with heart rate.

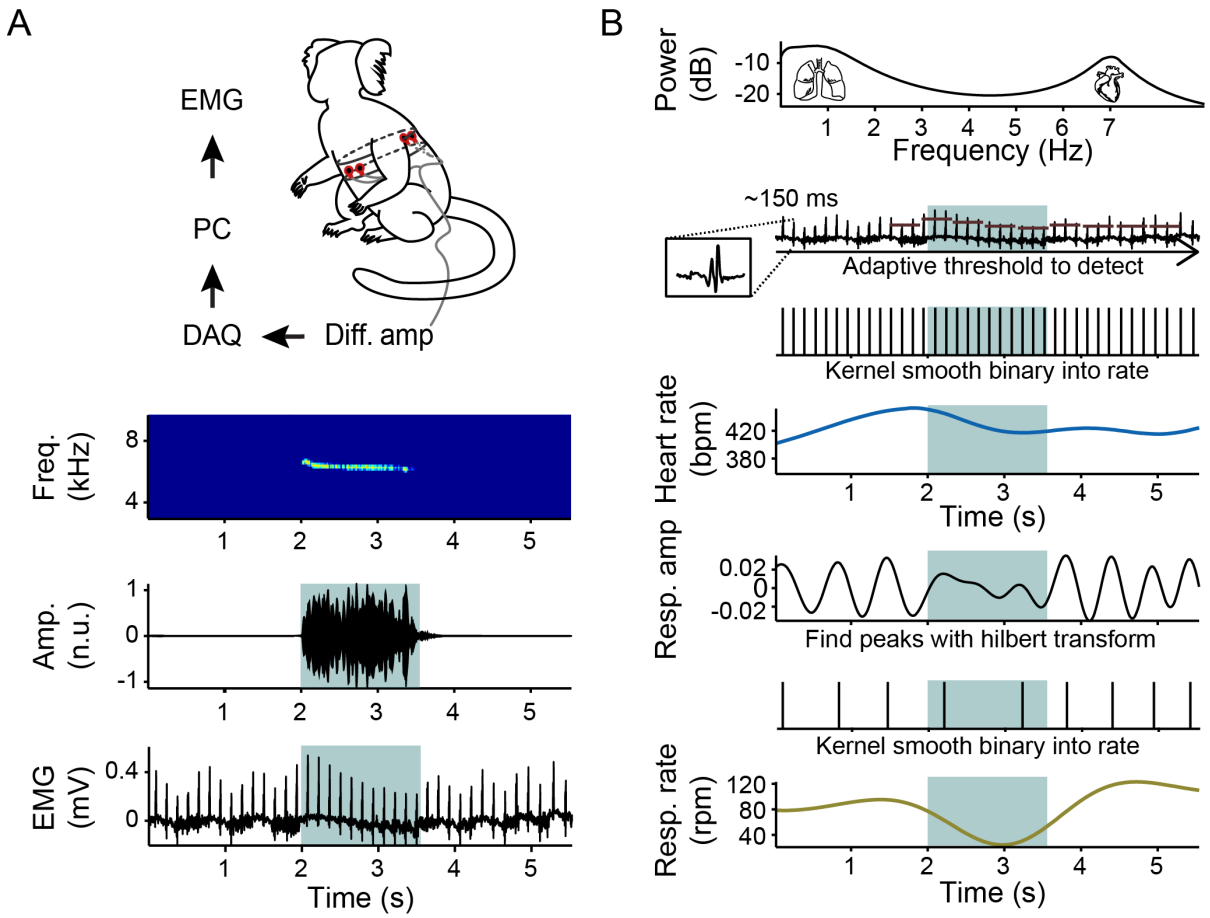


Figure 2.1 *Measuring respiration and heart rate from electromyographic (EMG) signal.* A: surface electrodes were applied to the dorsal and ventral thorax of an adult marmoset. Signals were processed through a differential amplifier (Diff. amp) before being sent to the data acquisition (DAQ) system (PC, personal computer). An individual exemplar of a vocalization (high-lighted in turquoise) with non-vocalization periods is plotted in black. The exemplar comprises the vocalization's spectrogram, the amplitude of the vocalization, and the raw data down-sampled to 1,500 Hz (Amp., amplitude; Freq., frequency; n.u., normalized unit). B: population power spectral density estimates demonstrating 2 peaks, one representing respiratory rate (0 –15 Hz) and another indicating heart rate (6 – 8 Hz). The cardiac signal output was isolated; an exemplar heart beat after filtering of the raw data is shown in the inset. Using an adaptive threshold, we detected the spikes exceeding the 95% percentile of the signal, converting the spikes into a binary 0 and convolving it with a 1-s Gaussian kernel, resulting in ongoing heart rate (bpm, Beats/min). After isolating the respiratory signal output through filtering, we calculated the angle of the Hilbert transform of the respiration signal whereby zero-crossing points represented respiratory peaks. We converted these peaks into a binary and convolved it with a 2-s Gaussian kernel, resulting in ongoing respiration rate (Resp. rate, respirations/min).

2.4.2 *The baroreceptor pathway does not wholly predict the link between heart rate changes and vocal production.*

We next examined the dynamics of heart rate during the moments preceding and following the production of a phee call. Phee vocalizations can differ in their durations (avg±sd, 2.03±0.96 seconds); we therefore plotted the heart rate data aligned to vocalization onset and offset, separately (Figure 2.3 A). Red lines indicate epochs that were significantly different from baseline cardiac activity (i.e., exceeded a 5% significance level). Marmosets exhibited an increase in heart rate prior to vocal production, followed by a rapid heart rate deceleration. The heart rate increase began slowly, starting roughly 23 seconds before the onset of a vocalization. At vocal offset, the steep deceleration in cardiac activity falls below the baseline for approximately 10 seconds. In Figure 2.3 A, the continuing increase in heart rate level corresponds to the increases in heart rate occurring with subsequent phee call production.

Considering that mammalian vocalizations are produced in a largely similar manner through the utilization of respiratory power and laryngeal tension (Fitch & Hauser, 1995; Ghazanfar & Rendall, 2008), any shared mechanisms across species are likely to include the baroreceptor pathways. The act of vocalizing is critically dependent upon respiration, and respiratory output is linked to cardiac activity via baroreceptor pathways (Critchley & Harrison, 2013). We wanted to test if the change in respiration during vocal production could directly account for the changes we observed in cardiac activity 23-seconds prior to phee call production (Figure 2.3 A). That is, maybe the vocal production-related cardiac changes we observed are not related to arousal per se, but are the consequence of respiratory sinus arrhythmia (RSA). RSA is a baroreceptor-driven activity arising from the interaction between the lungs, the nucleus of the solitary tract (NTS), and the heart. A signal is transmitted from the lungs via the vagus nerve to the NTS, where a reciprocal signal is sent

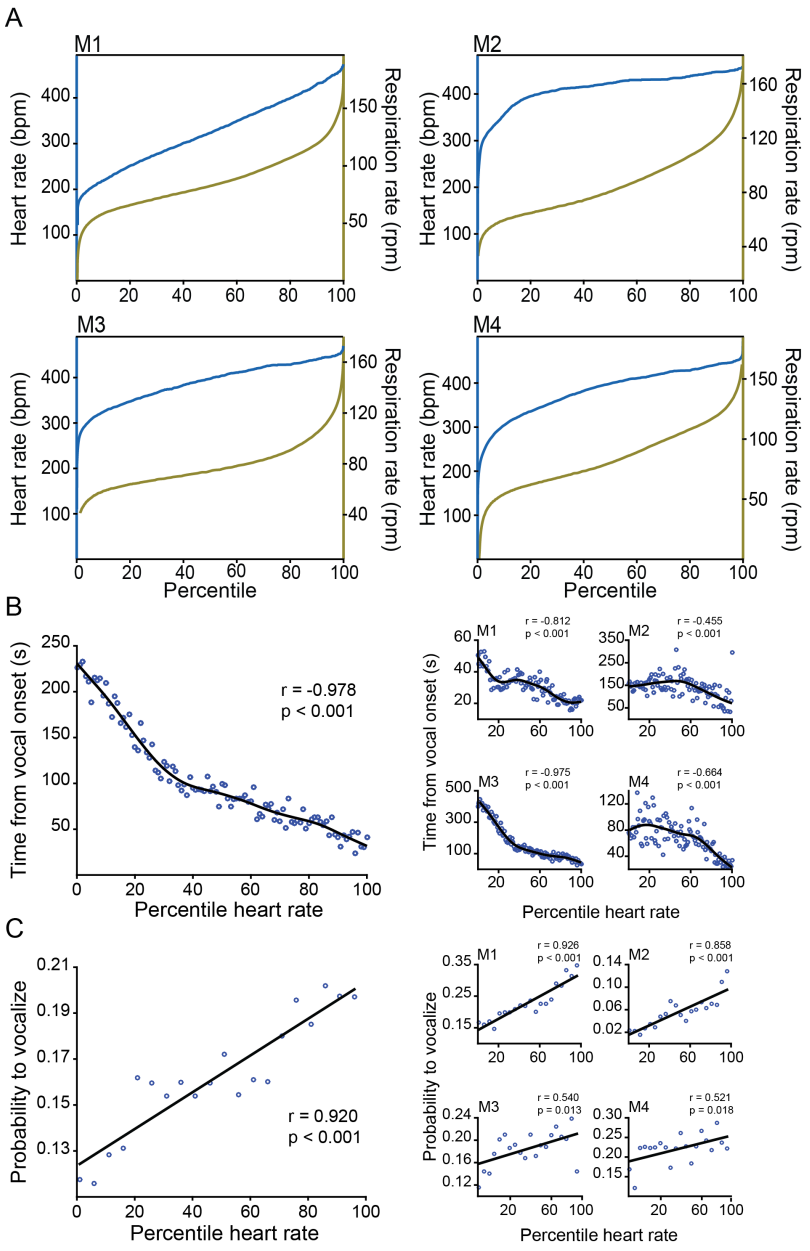


Figure 2.2 *Relationship between heart rate, respiration rate, and the production of vocalizations.* A: cumulative density plots for each animal (M1–M4) plotting heart rate (left y- axis, blue) and respiratory rate (right y-axis, green) against its percentile. B: average time from vocal onset plotted against percentile heart rate for the population with a cubic spline fit. At right, individual averages demonstrate the same positive correlation between time from vocal onset and heart rate percentile. C: average population probability of vocalization plotted against heart rate percentile. At right, probability curves are shown for each individual animal. Statistics reported are Spearman correlations.

back from the nucleus ambiguus to the heart, changing its dynamics (Stauss, Persson, Johnson, & Kregel, 1997): Inspiration (expiration) should increase (decrease) the heart rate. If respiration decreases, heart rate should increase, and vice versa. Considering this activity, the cardiorespiratory dynamics during vocal production could simply be the result of the baroreflex maintaining compensatory cardiac activity after the subject modifies its breathing to produce a vocalization. Figure 2.3 B shows that this compensatory mechanism indeed occurs during vocal production when respiration rate significantly (red-lined epochs) decreases while the heart rate increases. Moreover, at the offset of the vocalization, when the heart rate is decelerating, the respiration rate rebounds. However, not everything can be explained by this compensatory mechanism: For several seconds prior to the vocalization, when heart rate is increasing, there is no change in the respiration rate.

To explore the extent of the RSA contributions to the cardiorespiratory patterns observed around a vocalization event, we tested if heart rate could be linearly predicted from respiratory activity when a vocalization is produced. We first cross-correlated heart and respiratory rate signals (Figure 2.3 C). Next, we built linear predictors of heart rate from ongoing respiratory rate using both the positive and negative peak of correlation (Figure 2.3 C). Then, we created a 95% confidence region of predicted heart rate for the data aligned to call onset and call offset (Figure 2.3 D). When comparing the median heart rate activity to the predicted linear changes in heart rate activity, both the acceleration prior to vocal onset and the deceleration at vocal offset violated this prediction (exceeded a 95% confidence region). Therefore, during vocal production, respiratory rate fails to predict cardiac dynamics. These results suggest that cardiac-related interoceptive processes during vocal production cannot solely be predicted by the homeostatic interactions between the heart and lungs. Could another indicator of arousal account for the observed changes in heart rate?

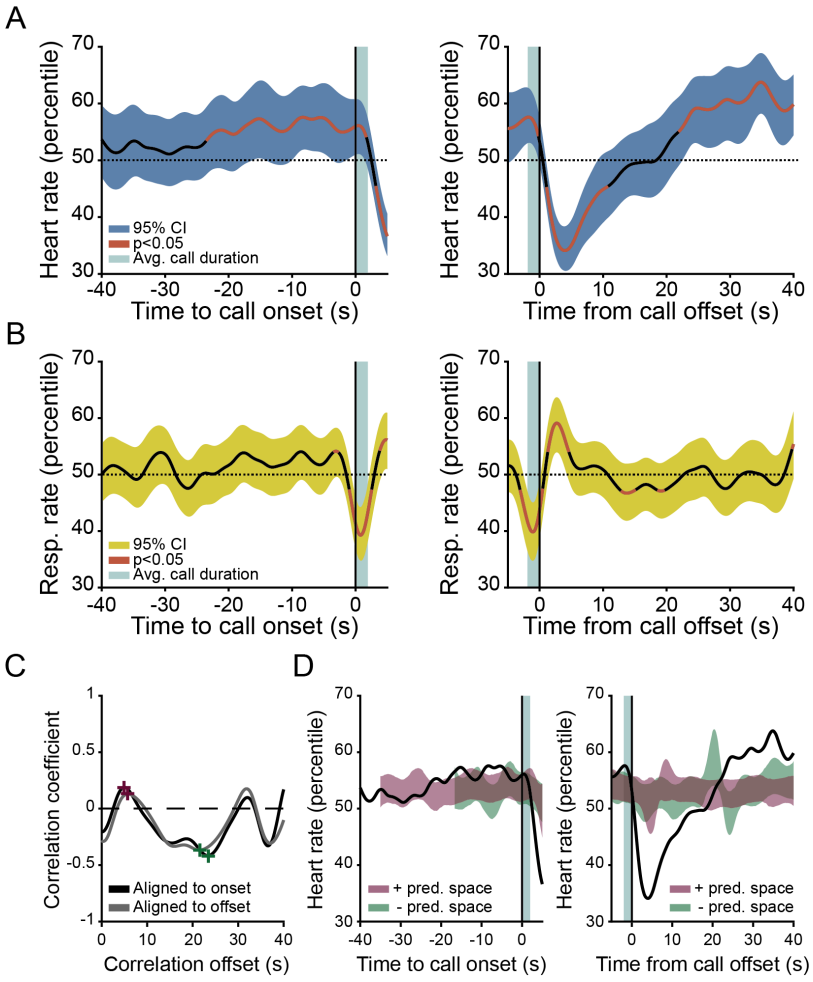


Figure 2.3 *Vocalizations are preceded by an increase in arousal.* A and B: smoothed population medians of ongoing heart rate (A) and respiration rate (B) plotted with the 95% bootstrapped confidence interval (CI) of the median. Red line indicates values outside 95% of the bootstrapped significance test for cardiorespiratory variability. C: cross-correlation plot of the respiration and heart rate signals for data aligned to call onset (black) and call offset (gray). The positive (purple) and negative (green) peaks of the cross correlation are denoted with crosses and were used to construct a linear predictor. D: results of a linear predictor of heart rate from ongoing respiration. The smoothed population median of heart rate is plotted (black) against the 95% bootstrapped CI of the respiration-predicted heart rate at its positive (purple) and negative (green) peak (Pred., predicted).

2.4.3 The link between heart rate changes and vocal production is not wholly predicted by general increases in activity levels

Across species, changes in general motor activity are considered to be a reliable indicator of arousal state (Garey et al., 2003; Pfaff, 2006). Physical activity is positively correlated with changes in heart rate, whereby cardiac activity is higher during movement than quiescence (Obrist et al., 1970). To determine whether changes in motor activity could predict the changes in heart rate observed during phee call production, we calculated a motion index for each animal using video recorded during the experimental session (Figure 2.4 A). Figure 2.4 B demonstrates the trace for the motion index 40 seconds before and after call onset. Red lines indicate epochs that were significantly different from baseline motor activity at the 5% significance level. Beginning 10 seconds prior to vocal onset, there is a gradual decrease in motor activity, with the cessation of movement reaching significance ~3 seconds before call onset. Once the call occurs, activity resumes. We sought to determine whether we could predict the observed cardiac dynamics using a linear predictor analogous to the one we built for cardiorespiratory activity (Figures 2.3 C, D). We constructed predictive regions of heart rate activity from the positive and negative peak of the cross correlation between motion and heart rate (Figure 2.4 C). Movement could predict the slow increase in cardiac activity, but not the deceleration after call offset (Figure 2.4 D). This analysis demonstrates that the increase in heart rate can be predicted by motor activity occurring 6-8 seconds before the change in dynamics. While predictive of heart rate dynamics prior to vocal onset, it is unclear whether the motor activity is causally related, as increases in cardiac activity during active movement typically occur instantaneously, usually within a cardiac cycle (Obrist, 1981; E. E. Smith, Guyton, Manning, & White, 1976).

Thus, neither respiration nor motion can wholly predict the fluctuations in cardiac activity. How, then, do patterns of cardiorespiratory activity correspond the spontaneous production of vocalizations?

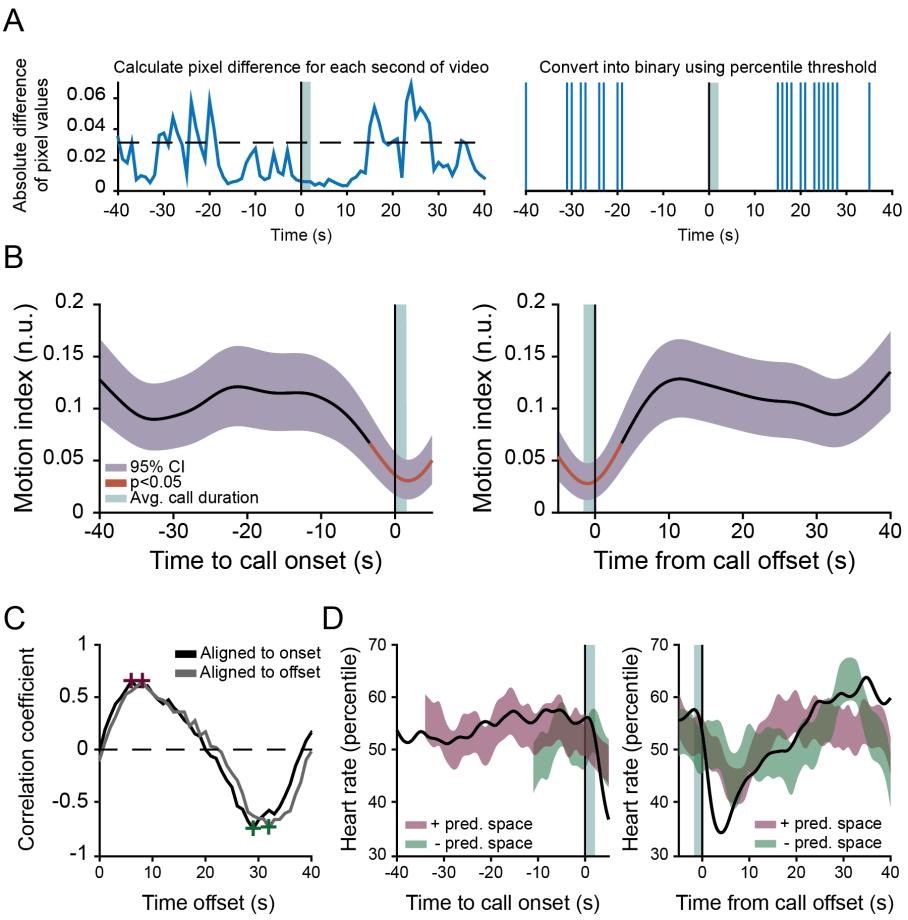


Figure 2.4 *General motor activity partially ac- counts for fluctuations in arousal during a vocalization*. A: schematic showing how we calculate the motion index for each animal. B: smoothed population median of ongoing motor activity plotted with the 95% bootstrapped CI of the median. Red line indicates values outside 95% of the bootstrapped significance test for motor variability. C: cross-correlation plot of the motion and heart rate signals for data aligned to call onset (black) and call offset (gray). The positive (purple) and negative (green) peaks of the cross correlation are denoted with crosses and were used to construct a linear predictor. D: results of a linear predictor of heart rate from ongoing motor activity. The smoothed population median of heart rate is plotted (black) against the 95% bootstrapped CI of the motion-predicted heart rate at its positive (purple) and negative (green) peak.

2.4.4 *Periodic autonomic nervous system fluctuations and vocal production: the Mayer wave.*

It is known that neural activity in the brain is modulated by rhythmic autonomic nervous system activity (Critchley & Harrison, 2013) and that the timing (phase) of external events relative to these interoceptive rhythms influence many behaviors (Barrett & Simmons, 2015). Such a relationship may exist between rhythmic autonomic nervous system activity and vocal production. We hypothesized that the drive to produce a vocalization, in the absence of a relevant and external conspecific signal, is phase-locked to rhythmic cardiorespiratory activity.

Figure 2.5 A shows an exemplar of a one-syllable phee call from Monkey 4 demonstrating a pronounced rhythmic oscillation in heart and respiratory rate. Figure 2.5 B is the probability density of time between the onsets of consecutive vocalizations and shows that spontaneous vocal production is rhythmic in nature (Takahashi et al. 2013). Figure 2.5 C shows an average power spectral density estimate of the period encompassing 40-seconds before a vocalization for heart rate (blue) and respiration rate (yellow). There is a peak frequency of 0.07 Hz for heart rate and 0.099 Hz for respiratory rate. These frequencies fall in the range of the Mayer wave, the resonant signature of the autonomic nervous system with a periodicity around 0.1 Hz (~0.04 - 0.15 Hz) and present in every mammalian species studied to date (Julien, 2006). Thus, in this context, the heart and respiratory rate rhythms are both products of the Mayer wave, the rhythmic activity of the sympathetic nervous system. We also observed a peak frequency of 0.145 Hz for phee call rate (Figure 2.5 C, in grey). To determine whether there was a relationship between the observed Mayer wave and call production, we calculated the phase of respiration and heart rate relative to the onset of phee calls. Figure 2.5 D reveals that the vocalization onset occurs after the peak of the Mayer wave, as measured by heart rate (blue, left), and just before the trough, as measured by respiratory rate (yellow, right). For both heart rate and respiration rate, these clusters around particular phases

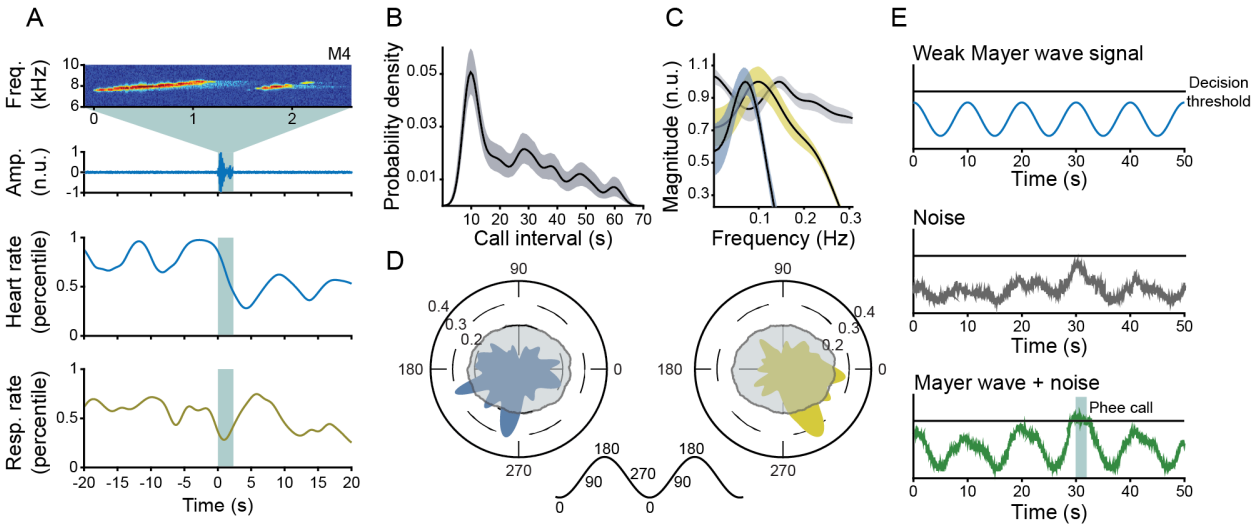


Figure 2.5 *Vocalizations are phase-locked to the rhythm of the autonomic system.* A: single exemplar of a 2-syllable vocalization The exemplar comprises the vocalization's spectrogram, the amplitude of the vocalization, the calculated heart rate, and the calculated respiratory rate. B: probability density showing the oscillatory nature of spontaneous phee call production. C: power spectrum density estimates for the call rate of marmoset vocal production (gray) and for the heart (blue) and respiratory rate (gold) preceding vocal onset. Shaded regions around averages represent 95% bootstrapped CIs. D: polar histograms of the phase location for the slow fluctuation in heart (left, blue) and respiratory rate (right, gold). Height corresponds to the percentage of values within the respective bin. Translucent histogram represents 95% bootstrapped significance level. At center, a schematic demonstrates the degree values against their respective locations on a schematic wave. E: a stochastic resonance model for how the Mayer wave influences spontaneous vocal production. The y-axis represents a neural activity.

exceeded the bootstrapped 95% confidence interval ($p<0.001$). The concentrations of the clusters were also significant (Cramer-Von Mises test). Therefore, the onset of spontaneous phee calls in the adult marmoset is preferentially located along a particular phase of the autonomic oscillation known as the Mayer wave.

Despite producing vocalizations in an oscillatory manner, marmosets do not produce a vocalization with every cycle of the Mayer wave. We hypothesize that this behavioral variability implies that the Mayer wave interacts with spontaneous vocal production in a manner similar to stochastic resonance (Figure 2.5 E). Under a model of stochastic resonance (Gammaitoni, Hänggi, Jung, & Marchesoni, 1998), vocal production occurs when a threshold is reached. The Mayer wave serves as a weak signal which, alone, is typically unable to reach the threshold and produce a call. If a source of noise adds upon the weak signal, the threshold is breached and vocal production occurs. In this model, the noise is unidentified, but could be one of the many processes that lead to a vocalization (see 2.5 Discussion).

2.4.5 *A dynamical systems view of vocal production*

Since the autonomic nervous system is critical for maintaining homeostatic control in the mammalian body (Loewy & Spyer, 1990), we sought to determine how the observed fluctuations in autonomic signals during vocal production correspond to baseline arousal state dynamics. To capture this baseline, we fit a separate multivariate Gaussian distribution to each pair of the bootstrapped dataset from heart and respiratory rate variability used in Figure3A-B and motor activity in Figure 2.4 B (Z-scored medians plotted in Figure 2.6 A). The circular regions in Figures 6B-D encompass the 95% confidence interval for each distribution. Against this baseline region, we plotted in grey the trajectory for the median of each dynamic for the entire population. Traces of heart rate, respiration rate, and motor activity begin 25 seconds before call onset, continue for the

duration of the call, and end 20 seconds after call onset. Figure 2.6 demonstrates the state space for the median arousal state for cardiorespiratory activity (Figure 2.6 B), cardio-motor activity (Figure 2.6 C), and respiratory-motor activity (Figure 2.6 D) around the production of a phee call. Green and red crosses denote the beginning and end of the trajectory, respectively, while the blue and black arrows signify call onset and the average call offset. The light blue line between the arrows delineates the call event. Surrounding the production of a phee call, we see numerous interactions between motor activity, heart rate, and respiratory rate. Cardiorespiratory activity oscillates at the edge of the 95% region of the state space, and ~23 second prior to the call, the activity leaves the baseline region as heart rate increases followed by a quick drop in respiratory rate ~ 2 seconds before call onset. This occurs as motor activity is gradually diminishing. During the call, motor activity has reached its trough and the direction of the dynamic switches. The dynamic activity traverses across the state space, returning to the baseline.

We consider this baseline as an attractor. A vocalization is produced when the dynamics are driven far away from this attractor. Thus, during the moments preceding and following a phee call, the arousal state is driven well outside the expected range of its variability. These trajectories, therefore, depict how interoceptive processes can influence spontaneous vocal production and the extent to which a vocal event, the phee call, propels the state beyond its typical activity range.

2.5 Discussion

We investigated the arousal dynamics related to the spontaneous, un-cued vocal production in marmoset monkeys. We exploited the fact that this species readily produces a single type of contact call - the phee call - within the undirected context of social isolation. This call is produced with some regularity (once every 10 seconds or so) without the need for an external sensory trigger such as the presence of, or vocalization from, a conspecific. We show that heart rate influences the timing and

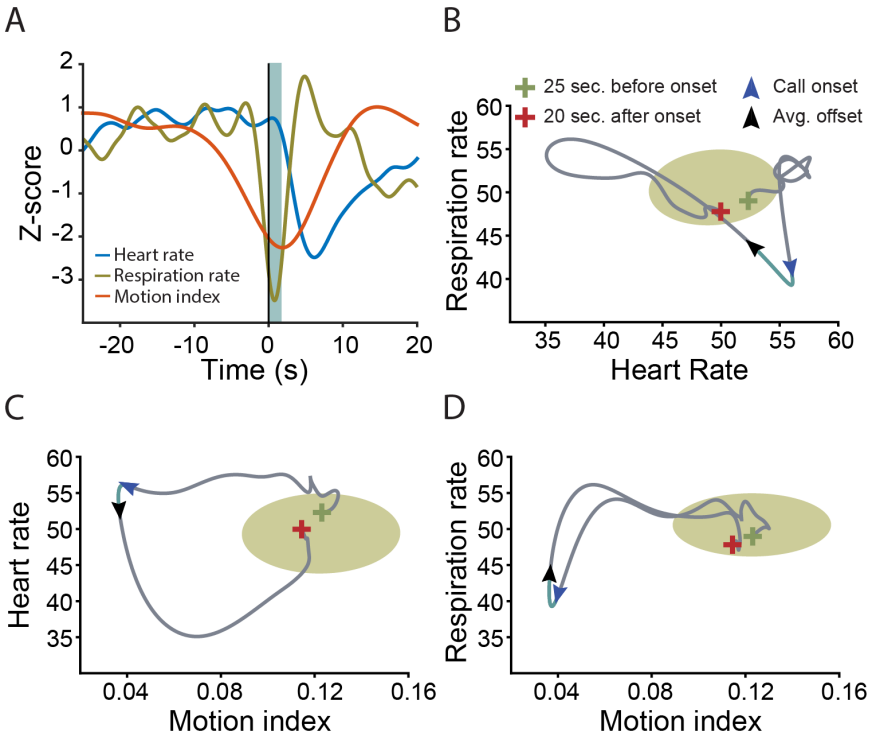


Figure 2.6 *Vocalizations are produced when perturbations of the system take the cardiorespiratory dynamics away from an attractor state.* A: Z- scored signals for heart rate (blue), respiration rate (gold), and motor activity (red) from 40 s before call onset until 20 s after call onset. Shaded region indicates the average duration of a call. B–D: state spaces for the median cardiorespiratory activity reported in Fig. 3, A and B. B: cardiomotor activity reported in Figs. 3A and 4B. C: and respiratory motor activity reported in Figs. 3B and 4B. D: Each graph represents median activity from 25 s before call onset until 20 s after call onset. The circular region outlines the 95% confidence intervals for multivariate normal Gaussian distribution of the bootstrapped significance test for the relevant data from Figs. 3A, 3B, and 4B. Green and red crosses denote the beginning and end of the trajectory, respectively, whereas the blue and black arrows signify call onset and the average call offset, respectively. The direction of the arrow indicates the dynamic's direction, and the light blue line between the arrows delineates the call event.

probability of vocal production. Gradual increases in cardiac activity are observed roughly 23 seconds prior to the onset of vocal production followed by a rapid and prolonged deceleration in heart rate of approximately 8 seconds in duration. The timing of vocal productions was related to the phase of on-going fluctuations of cardiorespiratory activity. These fluctuations were on the order of 0.1 Hz and represented the Mayer wave, a well-established indicator of sympathetic nervous system activity (Julien, 2006). It is important to note that the phee call is only one of several different calls produced by marmosets, but other types are heard in different contexts (Bezerra & Souto, 2008); our pattern of results may not generalize to these other call types.

The changes in heart rate that bracket the production of phee calls by marmoset monkeys are similar to those observed during human speech. Humans also exhibit an increase in heart rate prior to speaking (Linden, 1987; Lynch, Thomas, & Long, 1980), even during signed communication by deaf individuals (Malinow, Lynch, Foreman, Friedmann, & Thomas, 1986). Interestingly, nonverbal utterances in humans do not fit this pattern. Mirthful laughter elicited during structured play does not demonstrate any difference in heart rate before or after a bout of laughter (Fry & Savin, 1988). The cardiac and respiratory systems are highly intertwined such that the simple act of breathing air into the lungs temporarily blocks parasympathetic influence on the heart, increasing heart rate (Berntson, Cacioppo, & Quigley, 1993). Conversely, exhaling restores parasympathetic influence on the heart, decreasing heart rate. This dynamic is known as respiratory sinus arrhythmia (RSA) (Bernardi, Porta, Gabutti, Spicuzza, & Sleight, 2001; Berntson et al., 1993), and there is a correlation between it and heart rate (Grossman & Taylor, 2007). Since producing a vocalization requires changes in respiration, we tested whether the heart rate fluctuations before and after vocal production were simply due to the RSA and could thus be linearly predicted by ongoing respiration. We found, however, that respiration rate was insufficient in predicting both the increase in heart rate and its deceleration surrounding the call event.

Changes in heart rate prior to vocal onset were predicted by general motor activity of the animal; however a causal relationship is not clear as there was a 6-8 second delay between motor movement and heart rate increase. Considering that active movement typically induces near instantaneous changes in heart rate (Obrist, 1981; E. E. Smith et al., 1976), it is unclear in the current study whether the motor activity is causally related to the heart rate changes we observed. The acceleration of heart rate and decrease in motor activity prior to vocal onset could be the result of anticipatory effort to produce a call. However, in human reaction time studies, there is an observed anticipatory deceleration in heart rate prior to task performance (Jennings, van der Molen, Somsen, & Terezis, 1990; Obrist, Webb, & Sutterer, 1969; Somsen, Jennings, & van der Molen, 2004). Functionally, a state of cardiac deceleration following the production of a contact call implies a state of reduced arousal. Thus, calling in this manner may be as an act of self-soothing (Gracanin, Bylsma, & Vingerhoets, 2014).

The oscillatory signature of the autonomic nervous system - the Mayer wave - is present in both cardiac and respiratory measures. The Mayer wave represents perturbations to the baroreflex (Julien, 2006), resulting in an unstable negative feedback control loop that generates self-sustained oscillations at its resonance frequency. Most mammals exhibit a range of ~0.04 - 0.15 Hz. We observed a 0.145 Hz frequency of vocal production across our marmoset subjects and demonstrated that vocal onset occurs at a preferred phase of the ongoing Mayer wave. Yet, a vocalization was not produced in every cycle of the Mayer wave and this suggests a mechanism akin to stochastic resonance (Figure 2.5 E). Stochastic resonance requires three basic ingredients: (i) a threshold; (ii) a weak coherent input such as a periodic signal; and (iii) a source of noise that is inherent in the system (Gammaitoni et al. 1998). In our scenario, the Mayer wave is the periodic signal that, with the addition of as yet undefined source(s) of input, leads to a breach in the decision threshold, producing a vocalization with some probability greater than zero. In scenarios involving more than

one individual, the input signal would include the facial and/or vocal expressions of conspecifics (Ghazanfar & Takahashi, 2014a). Considering the relationship between vocal onset and the ongoing phase of the Mayer wave, an indirect consequence of this slow, pervasive autonomic oscillation is to assist in the assembly of vocal behavior. Indeed, vocal developmental data from marmosets suggests that the proper assembly of contact calling occurs postnatally, with an upward shift of the threshold (Y. S. Zhang & Ghazanfar, 2016). Infant marmosets in the undirected context produce vocalizations at almost 10 times the rate of adults, with a range of immature-sounding call types. After about a month, they slow the rate of call production and produce phee calls almost exclusively. The speed with which this transition from immature to mature phee calls takes place is influenced by contingent vocal feedback provided by parents (Takahashi et al., 2015).

We compiled the state space in which the combination of specific arousal states contribute to the production of a vocalization. The take-home message of that analysis, however, is not that arousal causes the production of a vocalization per se, but that perturbations to the resting state space increase the likelihood of a call occurring. Such perturbations are in the context of many other "moving parts" of the system that are operating on different timescales. For example, hormone levels influence the production of vocalizations and vice versa. Increased cortisol increases alarm calling in macaque monkeys (Bercovitch, Hauser, & Jones, 1995). Although there is no link between cortisol levels and the phee calls produced by marmosets in the undirected context (Norcross & Newman, 1999), the production of vocalizations decreases cortisol levels in this species (Clara, Tommasi, & Rogers, 2008; Cross & Rogers, 2004). Producing a vocalization is also metabolically costly, and energy levels are another factor that can influence vocal production (Ryan, 1988). Moreover, simply hearing the call of another can send the cardiorespiratory state space into fluctuation (Boysen & Berntson, 1989), and these changes in arousal level may influence the likelihood of a vocal response and the acoustic structure of this response (Choi et al., 2015;

Takahashi et al., 2013). Thus, the production of a vocalization is influenced by many factors: those occurring at slower time scales (e.g., hormones and energy levels) interacting with those occurring more quickly (e.g., sensory inputs, cardiorespiratory changes, neural activity).

In the undirected context studied here, the perturbations of resting arousal state were represented by three consecutive events: initial slow increase in heart rate followed by the cessation of movement, which is then succeeded by a quick decrease in respiration rate. Each of these three events drives the dynamics of the arousal system far away from its baseline region, resulting in vocal production. In dynamical systems terminology, the baseline region is an attractor and the vocalization is produced when a perturbation takes the dynamics away from the attractor. In this scenario, we hypothesize that the vocalization is a homeostatic mechanism to return the arousal dynamics to the basin of attraction, and finally into the attractor region. Thus, one could consider vocal production itself as a form of allostatic control (Berntson & Cacioppo, 2007; McEwen & Wingfield, 2003) or one parameter of a cybernetic system whereby compensatory actions (vocal production, in our case) offset the effects of disturbances (drifts away from the resting arousal state) (Pellis & Bell, 2011).

The neural pathways for these putative interactions involving arousal and vocal production overlap to a surprising degree and provide a viable pathway by which arousal influences vocal production. At the cortical level, the interoceptive system consists of a network of areas along the medial wall of the cerebral cortex, from the cingulate cortex to the ventromedial prefrontal cortex, as well as laterally into the orbitofrontal cortex and insula (Critchley & Harrison, 2013). These visceromotor cortices generate autonomic, hormonal and immunological predictions to adjust how the internal systems of the body deploy resources to deal with the external world in the immediate future (Barrett & Simmons, 2015). Collectively, they project subcortically to the amygdala, ventral striatum, hypothalamus and the periaqueductal grey (An, Bandler, Ongür, & Price, 1998; Critchley &

Harrison, 2013). This network organization overlaps with those related to primate vocal communication. The pathway from the anterior cingulate cortex to periaqueductal gray is critical for vocal production (Jürgens, 2009), and the insula contains a population of neurons that are exquisitely selective to conspecific vocalizations (Remedios, Logothetis, & Kayser, 2009). Subcortically, a key structure involved in regulating arousal (as measured by heart rate) is also critical for vocal production: the nucleus ambiguus of the reticular formation (Farkas, Jansen, & Loewy, 1997; Jürgens, 2009). Thus, on a neuroanatomical level, the neural systems governing vocal production are inextricably linked with the autonomic nervous system. Our immediate next steps are to link the cardiorespiratory dynamics described in the current study to the neurophysiology of vocal production.

Chapter 3

Vocal Behavior in Humans and Marmosets: Shared Development

3.1 Introduction

Many, if not all, behaviors are the outcome of dynamic developmental trajectories spanning multiple networks through time (Byrge, Sporns, & Smith, 2014). Besides sharing neural homologues and a suite of prosocial behaviors with humans (Borjon & Ghazanfar, 2014), the marmoset monkey exhibits markedly similar developmental trajectories for vocal production (Ghazanfar, Takahashi, Zhang, & Borjon, 2017). As adults, the timing of marmoset spontaneous vocal behavior is in sync with pervasive rhythms in the autonomic nervous system, demonstrating the inextricability of behavior and physiology (Borjon, Takahashi, Cervantes, & Ghazanfar, 2016). During development, behavior occurs within the constraints of a rapidly growing and changing body. The exact relationship between vocal development and the rest of the developing body is currently unknown. Here, we will review what we have learned from marmoset monkey vocal development as it relates to human vocal development. We will conclude by examining the putative mechanisms that may be operating across species.

3.3 Physiological influences on vocal development

Human vocal development is thought to be marked by progression. Early vocalizations, such as cries, fussing and cooing sounds, become increasingly complex and speech-like, eventually turning into words (Oller, 2000). From a purely acoustic perspective, the progression of changes from cries to later vocalizations like cooing and babbling is continuous (Kent & Murray, 1982). At around six months, infants begin to produce consonant-vowel combinations, signaling the transition into the

babbling period when their vocalizations most resemble speech. Early vocalizations have thus been considered to act as scaffolding for later, more speech-like vocalizations. This line of thinking, however, may be incompatible when we consider the behavioral dynamics of development. For an infant to produce a vocalization, it is necessary to successfully coordinate a rapidly changing body.

There is evidence that the development of vocal production follows milestones of motor development. Human infants naturally exhibit rhythmic facial movements early in development which are then used as a scaffold for new vocalization types (MacNeilage & Davis, 2000; MacNeilage, Davis, & Kinney, 1999). Subsequently, the emergence of rhythmic babbling, putative scaffold for mature speech, coincides with an increase in rhythmic arm movements (Thelen, 1979). Further, the onset of unsupported sitting in infants, a mature form of postural change, facilitates changes in the respiration of an infant. These respiratory changes bring about longer utterances with different types of vocalizations (Iverson, 2010). Physiologically, early infant vocalizations are also associated with an individual's state of arousal, a consequence of the autonomic nervous system (Zeskind, 2013). One measurable impact of arousal on vocal behavior in human infants includes predictable changes to the acoustic features of newborn cries (Stewart et al., 2013). The scaffolding of vocal production onto existing motor and autonomic activity is thought to be a critical factor that drives human vocal development (Ghazanfar et al., 2017; Ghazanfar & Zhang, 2016; Iverson, 2010; Iverson & Fagan, 2004). Supporting this idea, recent studies have reported a strong association between independent walking and language development, whereby infants who walk show larger vocabularies than those who crawl (He, Walle, & Campos, 2015; Walle & Campos, 2014).

At three months of age, human infants who successfully employ their ability to reach increase their visual attention to faces over objects (Libertus & Needham, 2011). Between five to seven months of age, human infants will be able to independently sit upright, without support, and

this behavior is associated with improved 3-D object completion (Soska, Adolph, & Johnson, 2010) That is not to say the development of behavior is a linear and modular succession: the onset of independent sitting also appears to temporarily disrupt an infant's ability to holistically perceive a face (Cashon, Ha, Allen, & Barna, 2013). The interplay between developing motor structures and vocal production is less well known in nonhuman primates. This is despite recent findings that describe marmoset vocal development to occur in a manner strikingly similar to humans. Below, we will review what is known about marmoset vocal development and will then continue to describe the development of the autonomic nervous system.

As in humans, variable, spontaneous vocal behaviors are ubiquitous in young marmoset infants (Takahashi et al., 2015). These early vocal behaviors primarily reflect the interplay between the infants' arousal states, sensorimotor coordination and biomechanical conditions (Y. S. Zhang & Ghazanfar, 2016). Contrary to what is typical for nonhuman primates (Egnor & Hauser, 2004), marmoset infants exhibit a dramatic change in vocal production in the first few months of postnatal life (Takahashi et al., 2015; Y. S. Zhang & Ghazanfar, 2016). At postnatal day 1 (P1), vocalizations produced are numerous and variable in their spectrotemporal structure. The number and acoustic variability of calls produced diminishes over 2 months, approaching mature vocal output whereby the infant marmoset will exclusively produce mature phee calls (Takahashi et al., 2013). Prior reports of much more subtle developmental changes in monkey vocalizations were frequently attributed to the passive consequences of growth in vocalization-related structures (Egnor & Hauser, 2004). However, the infant's growth rate, as indicated by changes in weight, does not accurately predict acoustic changes in vocal output (Takahashi et al., 2015). What, then could be driving this development? Before we consider the external drivers of vocal development, we should consider the prenatal factors that facilitate development in general.

Like any environment, the womb is a dynamic source of sensory stimulation (Harshaw & Lickliter, 2011). In this developmental niche, embryos are exposed to a variety of sensory signals, from parental vocalizations to gustatory input to the constant auditory stimulation of the mother's heartbeat and viscera. This prenatal sensory input provides a natural prelude to postnatal behavior. For example, changes in heart rate reveal the human fetus is sensitive to its mother's voice when compared to a stranger's voice (Lecanuet, Fifer, Krasnegor, & Smotherman, 2013); suggesting the fetus can perceive and learn auditorily. Prenatal factors interact with each other during the course of individual development (Gottlieb, 1992), by shaping postnatal behavior through influencing both an infant's arousal patterns as well as their biomechanics. For instance, in rodents, levels of arousal are linked to the actions of the estrogen receptor (Garey et al., 2003), and patterns of estrogen receptor expression are influenced by the epigenetics of maternal care (Champagne & Curley, 2008). Similar data are not available for marmoset monkeys, but other studies in this species demonstrate that early pre- and post-natal environmental events, including quality of parental care, influence patterns of infant behavior (Pryce, Aubert, Maier, Pearce, P. C., & Fuchs, 2011), including vocalizations (Dettling, Feldon, & Pryce, 2002) and gene expression in the brain (Law, Pei, Feldon, Pryce, & Harrison, 2009).

Moreover, changes in maternal arousal levels can influence the offspring both in the womb and through postnatal physical contact. For example, the cardiorespiratory dynamics of human infants will entrain to their mother's dynamics while laying on her body (Van Puyvelde et al., 2015) and during face-to-face communication involving vocalizations (Feldman, Magori-Cohen, Galili, Singer, & Louzoun, 2011). This maternal effect on human infants is present up to 2-3 months of age and is thought to be a continuation of similar prenatal maternal influences. Indeed, very young infants have difficulty self-regulating arousal levels; that regulation comes from parents (Moriceau & Sullivan, 2005). Similarly, in marmoset twins, common contact with parents via carrying and vocal

interactions (Takahashi et al., 2015) during the first week of postnatal life may influence their arousal levels simultaneously to generate similar patterns of fluctuations. In the following section I will explore the anatomical source of these arousal fluctuations in greater detail as well as their demonstrable influence on vocal behavior.

3.3 Social influences on vocal development

In humans, we know that social feedback from caregivers is an instrumental driving force that increases the maturation rate of prelinguistic vocalizations (Gros-Louis, West, & King, 2010). Parental eye contact, contingency, and responsiveness directly impact the quality and vocal pattern of infant vocalizations (Hsu & Fogel, 2001). Not only is the volubility of infants influenced by social context and past dyadic interactions with caregivers (Franklin et al., 2013; Goldstein, Bornstein, Schwade, Baldwin, & Brandstadter, 2009), but caregivers who preferentially and contingently respond to speech-like infant vocalizations spur the development of more complex vocalizations from those infants (Goldstein & Schwade, 2008; Goldstein, King, & West, 2003). Conversely, parents are more likely to contingently respond to infant vocalizations if they sound more mature or speech-like (Gros-Louis, West, Goldstein, & King, 2006).

Along similar lines, marmoset infants increasingly produce more mature-sounding calls subsequent to their parents' vocalizations over the course of development (Takahashi, Fenley, & Ghazanfar, 2016). Further, contingent vocal responses from marmoset parents influence the timing of the transition from immature to mature vocalization by reinforcing the production of mature phee calls (Takahashi et al., 2015). This is yet another parallel with human vocal development and consistent with the many ways infants can learn from parents beyond imitation (Tchernichovski & Marcus, 2014). In addition to facilitating the development of immature utterances into more mature sounding forms, such an interaction reinforces the development of vocal exchanges.

Cooperative vocal behavior is a ubiquitous form of social interaction in our lives. In humans, vocal cooperation develops during the first year of life (Elias & Broerse, 1996; Hilbrink, Gattis, & Levinson, 2015; Jasnow & Feldstein, 1986; Kajikawa, Amano, & Kondo, 2004; Stern, Jaffe, Beebe, & Bennett, 1975). In the first postnatal months, infants frequently overlap their vocalizations with their parent's utterances. By approximately 9 months of age, they can engage in proto-conversations with adult-like timing dynamics. In adult marmosets, this vocal behavior has the same universal features of human conversation, albeit on a different timescale, and exhibits the essential features of a coupled oscillator system (Takahashi et al., 2013).

Early in postnatal life, however, marmoset infants overlap their vocalizations with their parents' about 25% of the time (Takahashi, Fenley, & Ghazanfar, 2016). This overlap probability is consistent with human 3-month olds (Hilbrink et al., 2015). By two months of age, the probability of overlap for infant marmosets decreases to about 15%, which is also consistent with the near age-equivalent human infant at 18-months of age (Hilbrink et al., 2015). Adult marmosets rarely overlap their vocalizations (Takahashi et al., 2013). Thus, the fact that vocal cooperation is still developing at 2 months of age in marmosets is analogous to the way that human infants at 18 months of age still exhibit immature vocal interaction dynamics. Further, the amount of observed vocalization overlap is larger than what is expected by chance early in postnatal life (Takahashi, Fenley, & Ghazanfar, 2016), similar to what has been observed in 13 to 15 week old human infants (Elias, Hayes, & Broerse, 1986). Thus, marmoset infants undergo the same developmental trajectory for cooperative vocal behavior as humans, and do so during the same life-history stage (Takahashi, Fenley, & Ghazanfar, 2016).

3.4 The development of arousal dynamics

All behavior occurs within the constraints of the body, which is a complex and dynamic system that changes over time (Chiel & Beer, 1997; L. B. Smith & Thelen, 2003). One fundamental but poorly understood drive of behavior in all mammals is arousal, a multidimensional process with many indices that give us a window into the activity of the autonomic nervous system (Pfaff, 2006). Any animal exhibiting a high level of arousal is more alert to sensory stimuli, more active, and more reactive (Pfaff, 2006). A subset of measurable indices of arousal include heart and respiratory rate, pupillary dilation, and movement patterns (Pfaff, 2006). To date the exact relationship between fluctuations in arousal and natural behaviors is unclear. This is due to the emphasis of prior research on examining only single indices of arousal in isolation (Ebitz & Platt, 2015; Ebitz, Pearson, & Platt, 2014), and failing to take into account the multifaceted nature of arousal and its components. Some components of autonomic arousal patterns have a lawful relation to one another (Morin & Viala, 2002; Obrist, Light, McCubbin, Hutcheson, & Hoffer, 1979; Potts, 2002), such as the yoked relation between increased motor activity and increased heart rate (Obrist et al., 1970). Other indices are less directly related, possessing different time constants, such as the relationship between galvanic skin response and heart rate (Tursky, Schwartz, & Crider, 1970), or are uncorrelated in a natural setting (Lacey, 1967). Therefore, maximal information can be gleaned from the simultaneous measurement of multiple distinct indices of arousal (Borjon et al., 2016; Pfaff, 2006; Wass, de Barbaro, & Clackson, 2015).

As described in the previous chapter, one clear illustration of the utility of taking a dynamic view of the autonomic state is in the case of the marmoset monkey. The spontaneous vocal behavior of adult marmoset monkeys occurs after the predictable synchronization of the animal's arousal system (Borjon et al., 2016). First, there is a slow and continuous increase in heart rate ~23 seconds before a vocalization followed by a decrease in general motor activity ~11 seconds before the vocalization and finally a respiratory rate quickly decelerates ~3 seconds before vocal onset. These

changes in arousal need not be cognitively orchestrated as there are naturally occurring ebbs and flows in any mammal's autonomic state. Heart rate, blood pressure, and respiratory rate increase and decrease in predictable rhythms; one of which is the Mayer wave, a pervasive 0.1Hz oscillation present in all mammalian species studied to date (Borjon et al., 2016). In fact, the onset of a vocalization in the adult marmoset occurs just after the peak of the heart rate Mayer wave. What, exactly, is this Mayer wave and how would behavior adapt onto it?

The mammalian heart is not a metronome (Cygankiewicz & Zareba, 2013; Reyes del Paso, Langewitz, Mulder, van Roon, & Duschek, 2013). Folded in the constant beating of the heart is an array of oscillations that represent a myriad of biological functions in the body (Julien, 2006; Reyes del Paso et al., 2013). The spectral profile of the heart is usually classified into three families of oscillatory components: high frequency (HF), low frequency (LF) and very-low frequency. The high frequency band of cardiac activity falls within 0.15 to 0.40 Hz and represents the influence of respiration on heart rate. The low frequency band, within which the Mayer wave lies, falls within 0.04 to 0.15 Hz and is thought to represent fluctuations in the regulation of vasomotor tone and blood pressure. Very-low frequency oscillations in heart rate occur at less than 0.04 Hz and represent thermoregulation and kidney function. Just as these oscillations represent various biological functions in the body, the exact location of their source is varied and controversial.

The high frequency band of cardiac activity has also been termed respiratory sinus arrhythmia and is generally believed to be mediated by the vagus nerve. The very act of breathing causes rhythmic changes in the pressure within the chest. This causes differences in regional blood flow and cardiac output, resulting in rhythmic oscillations of blood pressure. The cardiac baroreceptor, a stretch-receptor located between the heart and lungs, senses these changes in blood pressure, venous outflow, and cardiac output and alters cardiac function to result in the high frequency oscillations we observe in cardiac activity. Selective administration of atropine in the vagus

nerve results in the elimination of high frequency cardiac oscillation. Therefore, these oscillations are likely the result of coupling between the vagus nerve, the baroreceptor, and respiratory drive. Developmentally, the magnitude of high frequency oscillations in heart rate tend to increase from the last trimester through the first postnatal months with stabilization around 2 years of age.

The origin of the low frequency band, and by extension the Mayer wave, is more controversial. One school of thought argues that low frequency oscillations of heart rate are indicative of sympathetic outflow and are generated by circuits in the brain stem, specifically the rostral and caudal ventrolateral medulla. Evidence is largely gathered from anesthetized cats (Montano et al., 1992), whereby sympathetic neurons exhibiting 0.12Hz activity were isolated and correlated with variability in heart rate and blood pressure. When the vagus nerve and sinoatrial node of the heart are resected, the activity in these sympathetic neurons persists. A second school of thought argues that low frequency oscillations are the consequence of interacting slow activation of the sympathetic nervous system and fast vagal responses. The baroreflex buffers slow, respiratory-induced oscillations in blood pressure and results in a resonant, low frequency oscillation. Behaviorally, we know that Mayer waves appear to influence behavior. In addition to their role in the spontaneous vocal behavior of marmoset adults (Borjon et al., 2016), these oscillations are present at birth (Y. S. Zhang & Ghazanfar, 2016), and that the acoustic structure of infant marmoset babbling oscillates at the same frequency of the Mayer wave.

Very-low frequency oscillations in cardiac activity are thought to represent kidney function and thermoregulation and are jointly controlled by both parasympathetic and sympathetic components of the autonomic nervous system. Developmentally, this is particularly interesting as thermoregulation changes dramatically postnatally, having tractable influences on behavior. Newborn mammals possess a reduced body size compared to adults. Small size is one of the most considerable factors influencing a newborn mammal's ability to regulate its own body temperature.

For example, newborn rodents are furless and unable to increase their body temperature through shivering (Blumberg, 2001; Leon, 1986). Without the presence of parents, littermates, or bedding, their body temperature will steadily decrease. To remedy this, groups of infant rodents will huddle together to maintain warmth, going so far as to demonstrate clear group regulation of temperature by increasing or decreasing the number of individuals in a huddle based on the ambient room temperature (J. R. Alberts, 1978; Harshaw & Alberts, 2012). On an individual level, the ultrasonic vocalizations of young rodents are highly sensitive to temperature (Blumberg & Alberts, 1997; Blumberg & Sokoloff, 2001; Harshaw & Alberts, 2012) whereby there is an increase in vocal production as the infant's body temperature decreases yet a diminishment in vocalization if their body temperature is kept stable through artificial means (Blumberg, Efimova, & Alberts, 1992b; 1992a). From the case study of the rodent, we propose a framework by which autonomic oscillations can become a scaffold for behavior to occur.

The behaviors exhibited by an infant animal are uniquely adapted for that life stage. For most infants, this includes a state in which thermoregulation is poor and core temperature needs to be maintained through huddling, swaddling, and interaction with others. We hypothesize, very-low frequency oscillations in cardiac activity will likely be visible and possess a high magnitude. As the infant grows, their ability to regulate and maintain their own temperature improves. This means that the developmental stage has changed. From this changed autonomic landscape, behaviors established earlier in development will adapt onto different internal dynamics of the autonomic nervous system. We know that the magnitude of low frequency oscillations in cardiac activity increase throughout development. Therefore, we hypothesize that a behavior previously exapted onto very-low frequency oscillations of heart rate would exapt onto the faster, low frequency, cardiac oscillations if the very-low frequency oscillations have diminished. In the case of the marmoset, we hypothesize that infant marmosets will exhibit vocal behavior in sync with very-low

frequency oscillations in cardiac activity. As their ability to thermoregulate increases, the behavior will begin to adapt onto low frequency oscillations and ultimately resemble the dynamics observed in the adult marmoset as described in our previously published work (Borjon et al., 2016).

3.5 Conclusion

Vocal signals are part of a complex, multidimensional, and probabilistic process that includes the physiological states of all the participants (Zeskind, 2013). The vocalizations of infant marmoset monkeys undergo dramatic changes during the first two months of life and their developmental trajectory closely mimics that of human infants. Parallel to these dramatic changes in vocal production, the marmoset itself grows, tripling its weight in a period of two months. As adults, marmoset monkeys demonstrate spontaneous vocal behavior that is in sync with a pervasive fluctuation in the autonomic nervous system, the Mayer wave. This oscillation is one of many different rhythmic signatures present in cardiac activity and in the autonomic nervous system. We propose that very-low frequency oscillations in cardiac activity, an index of thermoregulatory state, serve as an initial scaffold for behavior to occur in very young marmosets. Hypothetically, as the marmoset ages and their ability to thermoregulate increases, the behavior previously scaffolded onto very-low frequency cardiac oscillations begins to exapt onto the Mayer wave, a low frequency cardiac rhythm.

Chapter 4

The Arousal Dynamics of Marmoset Vocal Development

4.1 Abstract

Vocal behavior emerges during development through multiple interacting and nested processes. Prior research in humans has argued that vocalizations scaffold onto rhythmic motor behavior while parallel research in infant marmoset monkeys has proposed that naturally occurring fluctuations in arousal subserve vocalization. Here, we sought to determine the developmental relationship between arousal, vocal, and motor behavior. We leveraged the rapid development and high volubility of the marmoset monkey to compile developmental profiles of vocal production, motor activity, and arousal during the first two months of life. Consistent with the human literature, we demonstrate that motor development is correlated to vocal development in marmosets. We next demonstrate that the development of arousal is correlated to both vocal and motor development. Using a variant of structural equation modeling, we reveal that these three indices develop sequentially: autonomic development drives the development of motor behavior which drives the development of vocal behavior in the common marmoset.

4.2 Introduction

Development is a dynamic process that is the result of ongoing, interactive and nested factors unfolding across multiple time scales (Thelen & Smith, 2007). As development occurs, emergent behavior is constrained by both environmental context (Thelen, 1979; 1981) and the physical capacities of a given life stage (Hoffman, Flory, & Alberts, 1999). To overcome these constraints, new behaviors are often scaffolded onto previously existing forms and dynamics. For example, it has

been observed that the motor development of human infants is often predictive of the emergence of behavior in other, putatively unrelated, domains (Libertus & Violi, 2016). Specifically, the development of human vocal behavior has roots in non-cognitive processes that include motor activity and the autonomic nervous system (Ghazanfar et al., 2017; Ghazanfar & Zhang, 2016).

During development, vocal behavior requires the successful coordination of a growing and changing motor and autonomic system. Human infants naturally exhibit rhythmic facial movements early in development, which they are thought to use as a scaffold to produce new vocalization types (MacNeilage & Davis, 2000; MacNeilage, Davis, & Kinney, 1999). Subsequently, the emergence of babbling, a rhythmic vocal behavior, coincides with an increase in rhythmic arm movements (Thelen, 1979). Physiologically, early infant vocalizations are associated with the infant's state of arousal, a function of the autonomic nervous system (Zeskind, 2013). One measurable impact of arousal on vocal behavior in human infants includes predictable changes to the acoustic features of newborn cries (Stewart et al., 2013). The scaffolding of vocal production onto existing motor and autonomic activity is thought to be one critical factor that drives human vocal development (Ghazanfar et al., 2017; Ghazanfar & Zhang, 2016). Supporting this idea, recent studies have reported a strong association between independent walking and language development, whereby infants who walk show larger vocabularies than those who crawl (He et al., 2015; Walle & Campos, 2014).

The prolonged developmental trajectories of humans make it difficult to compile a complete profile of infant development through the simultaneous measurement of motor, vocal, and autonomic activity in the same individuals through time. Animal models, however, afford researchers the ability to more densely sample a profile of vocal development trajectories across multiple indices (Y. S. Zhang & Ghazanfar, 2016). Recent studies of infant marmoset vocal development have demonstrated compelling similarities to the vocal development of human infants

(Takahashi et al., 2015). Marmoset infants increasingly produce more mature-sounding calls subsequent to their parents' vocalizations over the course of development (Takahashi, Fenley, & Ghazanfar, 2016). Further, contingent vocal responses from marmoset parents influence the timing of the transition from immature to mature vocalization by reinforcing the production of mature phee calls (Takahashi et al., 2015). Here, we leverage the rapid development of marmoset infants to compile the first "whole-body" view of vocal development from birth until two months of age.

4.3 Materials and Methods

The subjects used in this study were seven infant common marmosets (Callithrix jacchus) bred at the laboratory from three different parental pairs (two sets of twins and one set of triplets, <2 months old). Both sets of twins and the triplets consisted of both genders. Each animal participated in surface electromyography recording during the first two months of life. The colony room was maintained at a temperature of approximately 27°C with 50%– 60% relative humidity, and a 12 hr light:12 hr dark cycle. Marmosets lived in family groups; all were born in captivity. All of them had ad libitum access to water and were supplied daily with standard commercial chow supplemented with fruits and vegetables. All experimental sessions were conducted during daylight hours between 12:00 and 18:00 and were performed in compliance with the guidelines of the Princeton University Institutional Animal Care and Use Committee.

4.3.1 *General experimental procedures*

The experimental setup was previously described in (Borjon et al., 2016; Y. S. Zhang & Ghazanfar, 2016), and Chapter 2 of this dissertation. We briefly summarize it here. Infant marmosets (n = 7) were separated from their parents and placed in a testing box in the experiment room. The triangular, prism shaped testing boxes were made of Plexiglas and wire (0.30 m x 0.30 m x 0.35 m).

Two pairs of the surface electrodes (Grass Technologies, silver-silver chloride) were attached to the infants, one pair on the chest close to the heart and one pair on the back close to the diaphragm. The front pair took the ECG signals and the back pair took the respiratory EMG signals. Electrode gel (Grass EC2) was applied on the surface of the electrodes to improve signal quality. Signals from each pair of electrodes were differentially amplified (gain = 250) and acquired at a sample rate of 40 kHz by the OmniPlex Neural Data Acquisition System (Plexon). We used the same heart rate analysis procedure previously described in (Borjon et al., 2016) and described in Chapter 2 of this dissertation.

Vocalizations were recorded using a Sennheiser MKH416-P48 microphone suspended 0.9m above the testing box. The microphone signal was sent to a Mackie 402-VLZ3 line mixer whose output was then relayed to the Plexon Omniplex for acquisition. This ensured the timing accuracy of the vocal signals relative to the EMG signals. We used the same criteria and algorithms established in previous work for computationally defining and segmenting infant marmoset vocalizations (Takahashi et al., 2015). This and all subsequent analyses were conducted in Matlab 2016b. Syllables were manually identified as phee, cry, or twitter using criteria established from prior research in this lab (Takahashi et al., 2015).

The 7 infants participated in a total of 220 surface electromyography sessions across the first two months of life (For Monkeys 1-7, respectively: 29, 29, 34, 34, 31, 31, 32). Each session lasted for 10 min. We recorded audio for 192 of the 220 sessions (For monkeys 1-7, respectively: 28, 29, 21, 20, 31, 31, 32) for a total of 10,956 calls (For monkeys 1-7, respectively: 1100, 801, 1604, 1021, 1657, 2199, 2574).

4.3.2 *Motor activity identification*

We recorded video for 215 of the 220 surface electromyography sessions (For Monkeys 1-7, respectively: 29, 28, 33, 33, 30, 30, 32) and identified a total of 3,195 instances of motor activity (For Monkeys 1-7, respectively: 662, 635, 414, 574, 207, 427, 276). We manually scored the recorded videos of each 10-minute experiment which were recorded at 30fps using Plexon Cineplex to identify the onset and offset of their motor activity using BORIS, an open-source event-logging software for video coding (Friard & Gamba, 2016).

Our definitions of motor activity were based on prior literature(Braun, Schultz-Darken, Schneider, Moore, & Emborg, 2015; Tardif, Layne, Cancino, & Smucny, 2002; Y. Wang, Fang, & Gong, 2014). Infant marmoset motor activity was subdivided into immature and mature behavior, then further subdivided into two categories: postural changes and locomotion. Postural changes were defined as instances where the infant repositioned itself and did not traverse across any distance in the testing box. Immature postural changes included raising of the head, forelimb support, hanging, and righting behavior. Mature postural changes included hindlimb support. The second category, locomotor movement, was defined as instances where the infant traversed the testing box. Immature locomotor movements included crawling and climbing. Mature locomotor movements include walking, jumping, or digging. As the video frame rate is at 30 Hz, the onset and offset of motor activity had a maximum resolution of 1/30 second.

4.3.3 *Statistical methods*

Unless otherwise stated, all p-vales are Holm-Bonferroni corrected. This statistical method corrects for multiple comparisons within the same dataset and mitigates the rate of false positives (Holm, 1979). All correlations are Spearman correlations, a statistical correlation method more resilient to outliers.

4.4 Results

4.4.1 *Marmoset infants undergo dramatic changes in vocal behavior and motor activity across development.*

The exact interactive nature between arousal, locomotor, and vocal development is currently unknown. One hypothesis is that these three processes develop independently and share little relation to one another (Figure 4.1A, top). A second hypothesis, derived from the human literature, is that motor development precedes vocal development. Prior literature has largely remained silent as to the role of arousal in guiding motor and vocal development (Figure 4.1 A, bottom). Below, we aim to address each of these hypotheses to determine the course of developmental processes occurring during the first two months of life.

The development of marmoset motor and vocal behavior is characterized by dramatic changes. Both postural and locomotor changes in motor activity have immature and mature forms (Braun et al., 2015; Tardif et al., 2002; Y. Wang et al., 2014). Together, changes in immature and mature postural and locomotor changes represent changes in general motor activity (Figure 4.1 B). Across development, the number of immature behaviors diminishes over time and the number of mature behaviors increases as the infants mature (Figure 4.1 C). Prior literature has extensively detailed the changes in vocal production during the first two months of life in the marmoset infant (Takahashi et al., 2015; 2016). One robust index of marmoset vocal development is a change in vocal entropy (Takahashi et al., 2015; Y. S. Zhang & Ghazanfar, 2016). Immature vocalizations are typically characterized with a high level of entropy while mature calls possess lower entropy (Figure 4.1 D). Across time, the average entropy of the calls produced by the marmoset decreases, acoustically resembling mature vocalizations (Figure 4.1 E). Consistent across both motor and vocal development, and a characteristic of any developmental process, is the decrease of immature

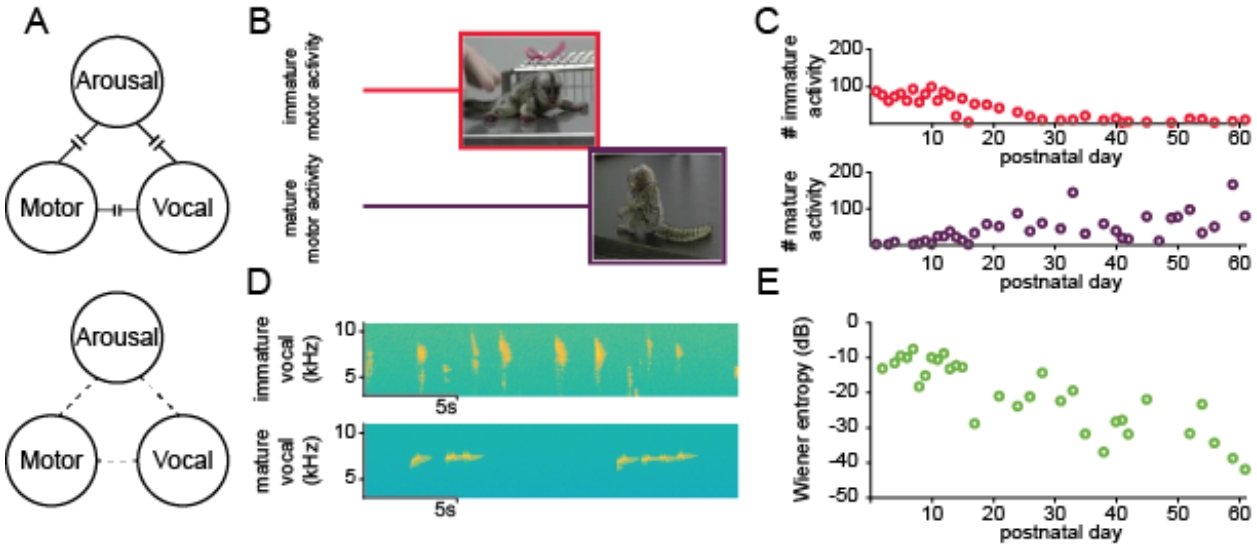


Figure 4.1 *The development of locomotor activity and vocal behavior in marmoset infants.* A: The hypotheses set forth in this chapter (top) the development of arousal, motor activity, and vocal behavior is uncorrelated and not connected. (bottom) Arousal, motor activity, and vocal behavior are related, but the directionality of their interactions is unclear. B: At top, images of immature and mature motor activity (top, raising head, bottom, hindlimb support, images taken from (Y. Wang et al., 2014)). C: The total numbers of immature and mature vocal behavior for each postnatal day recorded. D: 20 second spectrograms demonstrating immature and mature vocalizations from a marmoset infant at P02 (top) and P38 (bottom). E. The average entropy of calls produced for all 7 individual marmosets calculated over development. Each circle represents a different postnatal day.

behaviors and the increase of mature forms. Do they have any relation to each other or do they represent parallel, but independent processes?

4.4.2 *Marmoset motor development is correlated with vocal development*

Prior research in humans has argued that vocal development scaffolds on motor activity. We sought to determine whether such a relationship exists in marmoset infants (Figure 4.2 A). Regardless of age, all marmoset infants exhibit some level of motor activity across development. In general, the duration of movements tends to decrease over development, therefore more mature motor behavior is characterized by shorter movements. Comparing the percentage of time the infants exhibit motor activity to the average vocal entropy of the vocalizations produced at the same time (Figure 4.2 B), results in a positive correlation ($r = 0.22$, $p = 0.03$). If we look at the percentage of time marmoset infants exhibit immature motor activity, there is a consistent decrease across development (Figure 4.2 C, red). Parallel to this decrease, there is a simultaneous increase in the percentage of time the infant marmosets exhibit mature motor activity (Figure 4.2 C, purple). The decrease in immature motor activity is positively correlated with vocal entropy (Figure 4.2D, left, $r = 0.51$, $p = 0.03$) while the increase in mature motor activity is negatively correlated with vocal entropy (Figure 4.2 D, right, $r = -0.34$, $p < 0.001$). Across development, decreases in immature posture (Figure 4.2 E, red) and locomotion (Figure 4.2 G, red) are consistently positively correlated with vocal entropy (posture – Figure 4.2 F, left: $r = 0.48$, $p < 0.0001$; locomotion – Figure 4.2 H, left: $r = 0.39$, $p < 0.0001$). Increases in mature posture (Figure 4.2 E, purple) and locomotion (Figure 4.2 G, purple) are negatively correlated with average vocal entropy (posture – Figure 4.2 F, right: $r = -0.26$, $p = 0.01$; locomotion – Figure 4.2 H, right: $r = -0.3$, $p < 0.01$). Now that we have established a compelling correlation between motor development, at all levels, and the entropy of vocalizations produced in

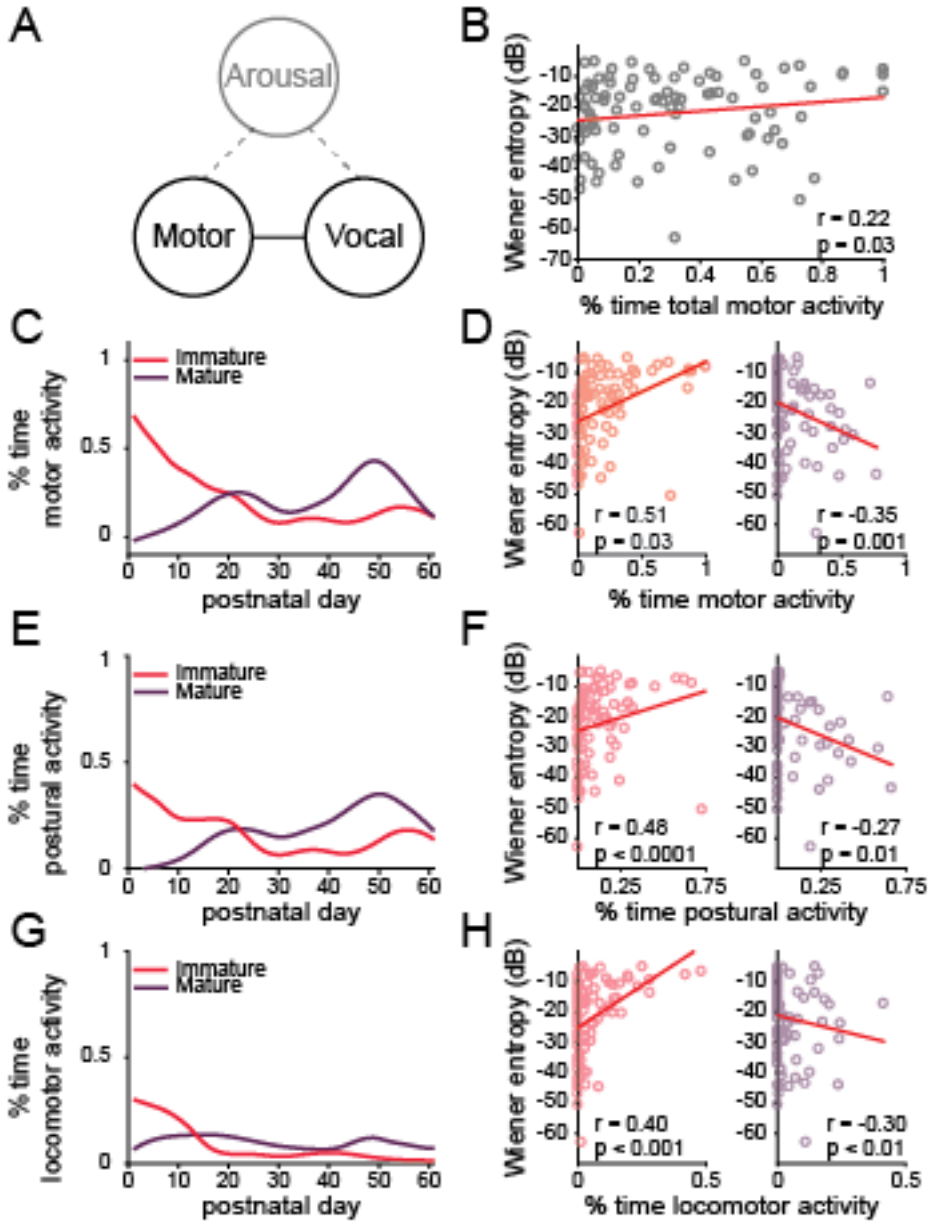


Figure 4.2 *Motor behavior is correlated to vocal behavior in marmoset infants.* A: The current state of our hypothesis: while directionality isn't known for marmoset infants, vocal behavior has been consistently associated with motor development in human infants. B: Spearman correlation for vocal entropy and the percentage of time an infant marmoset exhibits mature locomotion. For this and all further correlations in this figure, each circle represents a different postnatal day. P-values are Holm-Bonferroni corrected for multiple comparisons. C: Across development, the percentage of time infant marmosets exhibit immature behaviors (red) decreases while mature behaviors (purple) increases. D. Spearman correlations and linear fit for the data separated by immature (left) and mature (right) motor activity. E: The same data demonstrated in C, except divided by category of motor activity. Across development, instances of immature posture (red) decrease while instances of mature posture (purple) increase. F: Spearman correlations for immature posture (at left) and mature posture (at right) for the data presented in panel E. G: Instances of immature (red) and mature (purple) locomotor activity across development. H: Spearman correlations and linear fits for immature (left) and mature (right) locomotor activity with the average entropy of calls produced.

infant marmosets through time, we want to understand whether the arousal state of the marmoset exhibits developmental change and whether those changes correlate to our indices of vocal and motor development.

4.4.3 *Marmosets exhibit dramatic changes in their arousal state across development and these changes are correlated to both ongoing vocal and motor development.*

Across the first two months of life, marmoset infants exhibit dramatic changes in their arousal state. Early in life, infant marmosets are unable to sustain their heart rate which tends to gradually decelerate over time (Figure 4.3 A, top). As the infant matures, heart rate stabilizes, no longer exhibiting a steady decrease (Figure 4.3 A, bottom). Indeed, if we calculate the slope of heart rate for each session through development, we reveal early heart rate to possess a negative slope and that slope increases as the infants develop, gradually approaching zero: stability (Figure 4.3 B). As postnatal days progress, the average heart rate of marmoset infants also increases, doubling from an average of 200 beats per minute during the first week of life up to 400 beats per minute at the end of the second month (Figure 4.3 C). Normalized heart rate variability, the standard deviation of heart rate divided by the mean of heart rate, decreases throughout development, eventually reaching a floor level of variability (Figure 4.3 D). If we calculate a power spectrum density estimate of heart rate to demonstrate the spectral frequencies of cardiac activity, we also see a developmental change (Figure 4.3 E). Oscillations in cardiac activity between 0.04 Hz and 0.15 Hz represent the Mayer wave, a resonant signature of the autonomic nervous system which is relevant to spontaneous vocal behavior in adult marmosets (Borjon et al., 2016; Julien, 2006). Faster oscillations between 0.15 Hz and 0.4 Hz represent the influence of respiration on cardiac activity. As the infant matures, the magnitude of oscillations in the Mayer wave increase while the magnitude of respiratory-induced oscillations in cardiac activity decrease. To quantify this change, we calculated the ratio of low to

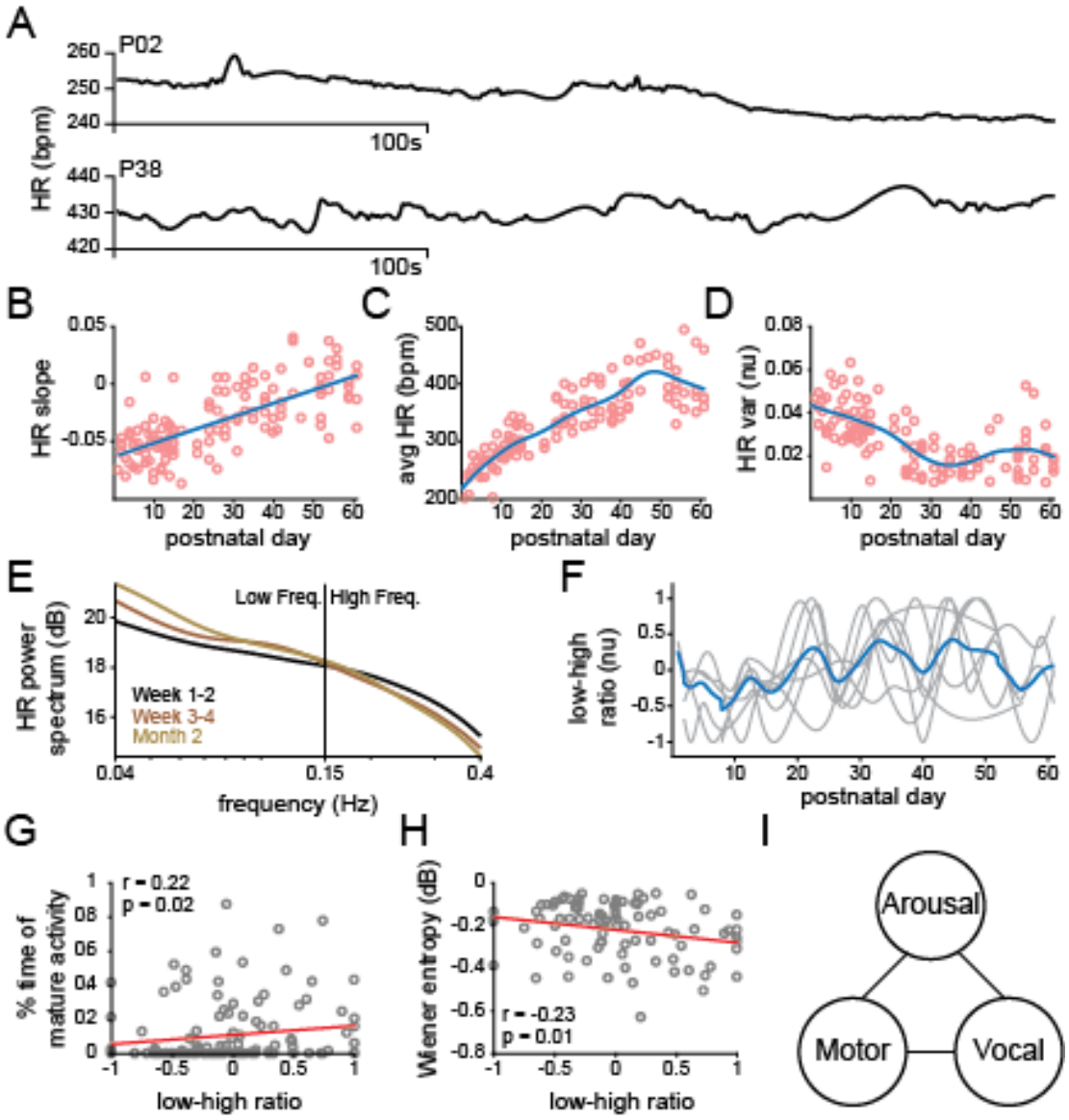


Figure 4.3 *Marmoset infants exhibit dramatic changes in arousal during development and these changes are correlated with vocal and motor development.* A: 300 second traces of immature cardiac activity collected at postnatal day 2 (top) and mature cardiac activity recorded t postnatal day 38 (bottom) from the same infant. B: The slope of heart rate activity for each infant at each postnatal day. Slopes were calculated by fitting a line to the entire 10-minute dataset gathered for each infant at each postnatal day. Linear fit (blue) represents the fit for the entire dataset across all infants. C: The average heart rate of the entire 10-minute session for each infant at each postnatal day. Smoothed average of the dataset in blue. D: The normalized standard deviation of heart rate for each infant at each postnatal day recorded. In blue, the smoothed average across all infants. E: Average power spectrum density estimates of heart rate for Weeks 1-2, Weeks 3-4, and the second month of life. F: Normalized developmental trajectory of the ratio between low and high frequency oscillations in cardiac activity. G: Spearman correlation with linear fit of the normalized low-high ratio against the percent of time an infant expresses mature motor activity. Each circle represents a different postnatal day for each infant. H: Spearman correlations for normalized arousal development against the average Wiener entropy. I. These results suggest direct relationships between the development of arousal, motor activity, and vocal behavior.

high frequencies (LH ratio) in cardiac oscillations for each day across development for each infant. A higher LH ratio indicates that Mayer wave oscillations are larger than respiratory-modulated oscillations. Figure 4.3 F demonstrates the normalized developmental trajectory of the LH ratio through time for each infant along with the average developmental trajectory. As captured by the LH ratio, autonomic development is, itself, oscillatory in nature.

We next asked whether this oscillatory pattern in arousal development is related to the vocal and motor indices collected. We plotted the normalized LH ratio against the percentage of time an infant marmoset exhibits mature motor behavior. If there is no relation between the development of motor activity and arousal, we would expect a flat line. Instead, we demonstrate a significant positive correlation (Figure 4.3 G, $r = 0.22$, $p = 0.01$). That is, when the LH ratio is above 0, and the strength of the Mayer wave is larger than that of respiratory influences on cardiac activity, the infant marmoset is exhibiting more mature motor activity.

We next correlated the normalized LH ratio to the average entropy of calls produced. There is a significant negative correlation between arousal development and the entropy of vocalizations produced (Figure 4.3 H, $r = -0.23$, $p = 0.02$). That is, infant marmosets exhibit vocalizations with more mature acoustic features when their LH ratio is higher. We therefore conclude that the development of motor activity, arousal, and vocal behavior are related in the infant marmoset (Figure 4.3 I). We next want to determine statistical causality between these three signals of development or whether these processes develop independently.

4.4.4 *Arousal development drives the development of motor activity, which drives vocal development.*

Development is a dynamic process that occurs over multiple time scales among many interacting factors (Byrge et al., 2014). When we consider the three developmental processes

of arousal, vocal behavior, and motor activity, it can be difficult to isolate a casual relationship as each process recruits multiple (and redundant) aspects of physiology. We therefore leveraged a pairwise linear non-Gaussian acyclic model (*Pairwise LiNGAM*) to statistically determine causal connections between our indices of development (Hyvärinen & Smith, 2013). This algorithm has been used in prior research to determine causality between nonlinear and non-Gaussian distributed variables with an unknown physiological origin (Faes, Porta, & Nollo, 2015; Kale et al., 2015) by computing the pairwise log-likelihood ratio amongst the loaded signals. By computing the connection strengths between the average vocal entropy, the percentage of time an infant exhibited mature motor activity, and the LH ratio, we demonstrate that the LH ratio drives motor development which then drives the average vocal entropy (Figure 4.4 A). The connection diagram reveals connections between the LH ratio and motor activity, motor activity and vocal behavior, as well as a third connection between the LH ratio and vocal behavior (Figure 4.4 C). Vocal behavior itself did not connect to either of the two other variables. To determine whether the connection strengths we observed were not due to chance, we constructed two bootstrap significance tests. The first test was to calculate the bounds of connection strengths by randomly permuting the LH ratio 1,000 times and comparing the 2.5^{th} and 97.5^{th} percentile of the bootstrapped connection strengths with the value observed. Both the connection between the LH ratio and motor activity, and the connection between the LH ratio and vocal behavior were significantly outside the bounds of what could be expected by chance (Figure 4.4 B). Our second significance test randomly bootstrapped the connection between motor activity and vocal behavior 1,000 times. The observed connection strength between motor

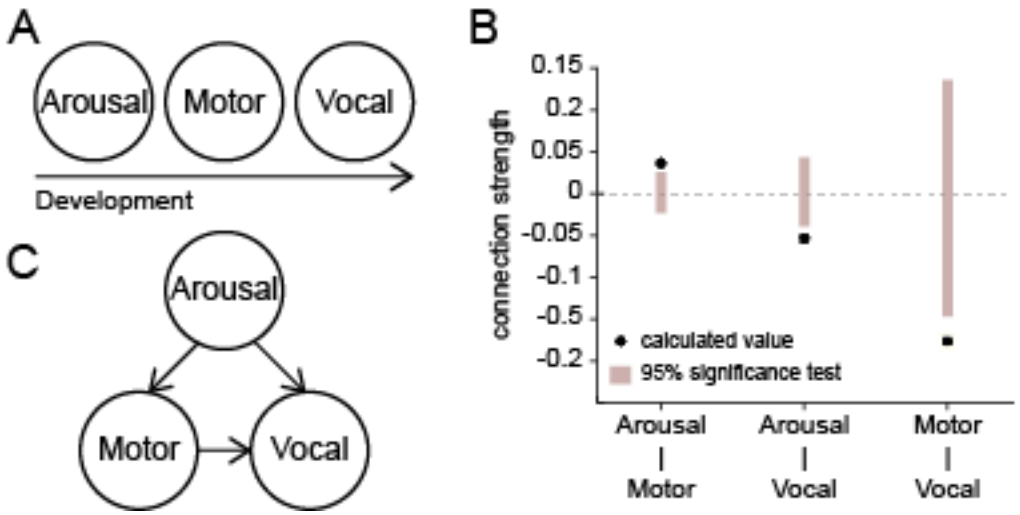


Figure 4.4 *Arousal development drives vocal behavior which drives the development of motor behavior.* A: Schematic indicating the developmental trajectory of the three indices analyzed. Pairwise LiNGAM analysis of causality demonstrate that arousal precedes the development of motor activity while motor activity drives vocal behavior. B: Bootstrapped significance test for the connection strengths calculated by Pairwise LiNGAM. Shaded regions encompass 95% of the permuted regions while black circles indicate the calculated values. C: Statistically significant connections exist between arousal and motor and vocal development while a statistically significant connection relates motor development with vocal behavior.

activity and vocal behavior was significantly outside the bounds of what could be expected by chance (Figure 4.4 B). We therefore demonstrate statistically causal connections between the development of the autonomic nervous system, as defined by the LH ratio, the development of motor activity, as quantified by the percentage of time an infant exhibits mature behavior, and the development of vocal behavior, as characterized by the average entropy of vocalizations produced. Together, these data argue that marmoset vocal development relies on changes in both the autonomic nervous system and motor activity.

4.5 Discussion

While much prior research has focused on the cognitive aspects of speech development, less attention has been paid to the motor milestones preceding, and often predicting, human speech development (Iverson, 2010). Parallel research in marmoset infants has argued that fluctuations in arousal, a function of the autonomic nervous system, are critical for the development of mature vocal behavior (Ghazanfar & Zhang, 2016; Y. S. Zhang & Ghazanfar, 2016). We sought to determine the interactive nature between developing motor systems, arousal dynamics, and vocal behavior.

We leveraged the rapid development and high volubility of the common marmoset to densely sample indices of motor and vocal development across the first two months of life. We categorized infant marmoset motor activity as immature or mature changes in posture or locomotion. Then we compiled the number of movements and the percentage of time an infant marmoset exhibited motor activity during the 10 minute sessions. Through development, marmoset infants exhibit less immature motor activity and increase the prevalence of mature motor behavior (Figure 4.1 C). One robust characteristic of marmoset vocal development is a decrease in entropy in the acoustic features

of their vocalizations (Figure 4.1 E). Mature marmoset calls are characterized by being low in entropy and this developmental trajectory is robust across individuals (Takahashi et al., 2015).

Prior research has demonstrated that human vocal development is underpinned by changes in motor function (Iverson, 2010; Iverson & Braddock, 2011; Libertus & Violi, 2016; Thelen, 1991). For example, preceding rhythmic babbling, human infants begin to exhibit rhythmic arm movements (Thelen, 1981a). Further, the onset of unsupported sitting in infants, a mature form of postural change, facilitates considerable changes to respiration and an infant's speech articulators and can bring about longer utterances characterized by different types of vocalizations (Iverson, 2010). In this study, we show that marmosets also display a relationship between motor and vocal behavior. We correlate the entropy of vocalizations produced by marmoset infants with the development of their motor activity. The less time a marmoset infant exhibits immature motor behaviors, the less entropy will be produced in their vocalizations, meaning their vocalizations become more mature (Figure 4.2 D, left). Similarly, the more time an infant exhibits mature motor behaviors, the more mature their vocalizations become (Figure 4.2 D, right). This relationship is robust across both postural (Figure 4.2 F) and locomotor (Figure 4.2 H) changes in development. Vocal behavior and motor activity are not the only processes undergoing dramatic changes in development: the infant marmoset is also experiencing changes in their internal, arousal state.

Like many small mammals (E. F. Adolph, 1971), the infant marmoset experiences drastic changes in physiology during its first two months of life. When separated from their parents, immature infant heart rate is characterized by a steady decline in heart rate which is likely representative of a lack of thermoregulatory capacity (Figure 4.3 A, top). Marmoset infants do not fully wean until three months of age and are usually carried by their parents and unrelated helpers during their first few months of life (Borjon & Ghazanfar, 2014). Later in development, after the first month of life, infant heart rate stabilizes and does not exhibit the same decline seen earlier

(Figure 4.3 A, bottom) which is likely indicative of the infant possessing the capacity to maintain its internal regulatory capacity, homeostasis. Parallel to this stabilizing of heart rate (Figure 4.3 B), there is an increase in average heart rate across development (Figure 4.3 C) as well as a decrease in heart rate variability (Figure 4.3 D). While human infants, and indeed many mammals, exhibit a general deceleration in heart rate over development, it is common for very small mammals to experience a general increase in average heart rate from infancy to adulthood (E. F. Adolph, 1971). Consistent with our marmoset data, human infants exhibit a decrease in heart rate variability through development (Richards & Cameron, 1989).

Over the first two months of life, we demonstrate an increase in the magnitude of low frequency oscillations in heart rate, representing the Mayer wave, and a decrease in high frequency oscillations (Figure 4.3 E) which are modulated by respiratory activity (Julien, 2006; Reyes del Paso et al., 2013). Prior work in adult marmosets has demonstrated that their natural and spontaneous vocal behavior is rhythmically locked to the phase of the Mayer wave (Borjon et al., 2016). The development of the LH ratio, our index of spectral change in heart rate, is itself oscillatory (Figure 4.3 F). We correlate the development of this ratio to changes in motor and vocal behavior and demonstrate that, on days where the magnitude of low-frequency oscillations is higher than high-frequency oscillations in cardiac activity, infant marmosets spend more time exhibiting mature motor activity (Figure 4.3 G) as well as less entropy in their vocalizations (Figure 4.3 H). We can therefore associate the oscillatory development of cardiac activity to both motor and vocal development in the infant marmoset.

To determine whether a causal order exists between the development of arousal, vocal, motor activity, we leveraged a causal inference algorithm. Specifically designed to calculate causality without the presence of known priors, this algorithm uses pairwise likelihood ratios to infer the causal direction between pairs of variables (Hyvärinen & Smith, 2013). We can then sequentially order the

statistically causal relationship between our observed developmental processes. The LH ratio drives the development of mature motor activity which then drives vocal development (Figure 4.4 A). In addition, there is a significantly strong connection strength between the LH ratio and vocal development (Figure 4.4 B:C). Therefore, we can consider arousal to be an important engine facilitating the development of mature behaviors. This notion is consistent with prior work in marmoset infants (Ghazanfar & Zhang, 2016) and has implications for human development (Ghazanfar et al., 2017).

In this study, we densely sampled multiple indices of development to compile a whole-body view of vocal development in the common marmoset. Consistent with the human literature, we demonstrate that motor activity drives the development of vocal production. We also demonstrate that the infant's autonomic nervous system undergoes dramatic changes in the first two month of life. Statistical measures of causality indicate that arousal serves as a scaffold for both mature vocal and motor behavior to occur.

Chapter 5

General Discussion

Humans and marmosets share similar characteristics of cooperative vocal behavior (Borjon & Ghazanfar, 2014; Takahashi et al., 2013), despite marked differences in brain size between the two species. This is one indication that certain aspects of human-like vocal behavior and communication may not require a large brain or complicated cognitive mechanisms (Borjon & Ghazanfar, 2014). An observation made in Chapter 1 was that the natural vocal exchanges between two marmosets exhibit an oscillatory pattern (Takahashi et al., 2013). Instead of leaning on a theory of communication that relies on the mutual exchange of information between a sender and receiver (Flack, 2013), or one that emphasizes the goal of communication (Rendall et al., 2009; Rendall & Owren, 2009), we argued that the oscillatory dynamics of the interaction itself belie the need for complex mechanisms or intentions.

Like the marmoset's vocal exchanges, it has been empirically demonstrated that their spontaneous vocal productions are also oscillatory (Borjon et al., 2016; Takahashi et al., 2013). If we imagine each marmoset as an oscillator, a coupling of two oscillators, two individuals, would result in the conversational dynamics observed in marmosets, and could serve as a putative evolutionary mechanism for human vocal cooperation. One potential oscillatory drive of vocal production is the animal's own state of arousal. Evolutionarily, there is reason to suspect that the modulation of arousal levels contributes to vocal behavior. It has been shown that auditory contact between marmosets through phee calls reduce cortisol levels (Rukstalis et al., 2003), while unmodulated arousal can be damaging and even fatal to primates (Uno et al., 1989). Thus, modulation of arousal levels and cooperative behaviors are closely intertwined (Rukstalis et al., 2003). Consistent with this

idea, it has been posited that human conversations play the same role as social grooming seen in apes and monkeys (Dunbar, 1998), suggesting that, in a sense, humans and marmosets have achieved "grooming at a distance" (Takahashi et al., 2013).

Throughout this dissertation, I have leveraged the volubility and rapid development of the common marmoset to examine how "non-brain" mechanisms, such as physiological measures of arousal, are predictive of vocal production and facilitate its development.

All behavior occurs within the constraints of the body, which is a complex and dynamic system that changes over time (Byrge et al., 2014). One fundamental but poorly understood drive of behavior in all mammals is arousal, a multidimensional process with many indices that give us a window into the activity of the autonomic nervous system (Pfaff, 2006). Measureable indicators of arousal state include heart and respiratory rate, pupillary dilation, and movement patterns (Loewy & Spyer, 1990; Pfaff, 2006). To date, the exact relationship between fluctuations in arousal and natural behaviors is unclear. Prior research has focused on examining only single indices in isolation (Ebitz et al., 2014; Ebitz & Platt, 2015), which disregards the multifaceted nature of arousal and its components. While some autonomic arousal patterns have a lawful relation to one another (Morin & Viala, 2002; Obrist et al., 1979; Potts, 2002), such as the yoked relation between increased motor activity and increased heart rate (Obrist et al., 1970), other indices occur at different time constants and are less directly related, such as the galvanic skin response and heart rate (Tursky et al., 1970), or are uncorrelated in a natural setting (Lacey, 1967). Therefore, maximal information can be gleaned from the simultaneous measurements of multiple distinct indices of arousal (Borjon et al., 2016; Pfaff, 2006). Throughout this dissertation, we have aimed to cover multiple indices of arousal by non-invasively monitoring heart rate, respiration rate, and motor activity in the marmoset, through infancy and adulthood, during naturalistic, untrained behaviors while in isolation.

In Chapter 2, we aimed to directly test whether our hypothesis about autonomic influences on vocal behavior could be verified empirically. We wanted to understand whether the natural, spontaneous vocal behavior of marmosets had any connection to ongoing arousal dynamics. In primates, prior literature has consistently claimed that changes in interoceptive state – i.e. arousal – govern primate vocalizations (Hammerschmidt & Fischer, 2008; Owren et al., 2011; Rendall & Owren, 2009). Yet we know very little about the nature of autonomic influences on the likelihood and timing of vocalizations. We thus investigated the arousal dynamics related to spontaneous vocal production in marmoset monkeys.

We exploited the fact that this species readily produces a single type of contact call during an undirected context of social isolation. The "phee" call is produced at roughly regular intervals of 10 seconds or so, even in the absence of an external sensory trigger such as the presence of, or vocalization from, a conspecific. In our study, we showed that heart rate influences the timing and probability of vocal production such that the timing of vocal productions was related to the phase of on-going fluctuations of cardiorespiratory activity. These fluctuations were on the order of 0.1 Hz and represented the Mayer wave, a well-established indicator of sympathetic nervous system activity (Julien, 2006). It is important to note, however, that the phee call is only one of several different calls produced by marmosets that are heard in different contexts (Bezerra & Souto, 2008); our pattern of results may not generalize to these other call types.

While the results from Chapter 2 exposed a compelling correlation, they fell short of a direct causal manipulation. A direct causal manipulation of the arousal state is difficult because pharmacological manipulations act on timescales that are difficult to predict and will last for varying half-lives (Bonyhay, Risk, & Freeman, 2013; Reyes del Paso et al., 2013; Stanaszek, Kellerman, Brogden, & Romankiewicz, 1983; van de Borne et al., 2001). We instead decided to determine whether causality

could be established by observing the development of marmoset vocal behavior. After all, many, if not all, behaviors are the outcome of dynamic developmental trajectories (Byrge et al., 2014).

Besides sharing neural homologues and a suite of prosocial behaviors with humans (Borjon & Ghazanfar, 2014), the marmoset monkey exhibits markedly similar developmental trajectories for vocal production (Ghazanfar et al., 2017). The vocalizations of infant marmoset monkeys undergo dramatic changes during the first two months of life and their developmental trajectory closely mimics that of human infants (Ghazanfar et al., 2017). Parallel to these dramatic changes in vocal production, the marmoset itself grows, tripling its weight within two months (M et al., 1992; Phillips, 1976; Schultz-Darken, Braun, & Emborg, 2015). While much prior research has focused on the cognitive aspects of speech development in humans, less attention has been paid to the motor milestones preceding, and often predicting, human speech. On the other hand, parallel research in marmoset infants has argued that fluctuations in arousal, a function of the autonomic nervous system, are critical for the developing of mature vocal behavior.

Human vocal development has been shown to be underpinned by changes in motor function. Rhythmic arm movements in human infants precedes rhythmic babbling (Thelen, 1981b). Further, the onset of unsupported sitting, a mature form of postural change, facilitates considerable changes to respiration and an infant's speech articulators, and can bring about longer utterances characteristic of different types of vocalizations (Iverson, 2010). Vocal development in marmoset infants reveals changes in the spectral entropy of vocalizations, such that calls with decreasing entropy more closely resemble the mature form (Takahashi et al., 2015). In Chapter 4, we demonstrated that motor activity in the developing marmoset correlates with vocal development. Both the decrease in immature motor behaviors and the concurrent increase in mature motor behaviors coincided with the production of more mature vocalizations. This relationship was robust across both postural and locomotor changes in motor development. Next, we sought to extend this understanding of

physiological influences on vocal development to the developmental changes of the autonomic nervous system and how it relates to our calculated indices of motor and vocal development.

The infant marmoset experiences drastic changes in physiology during its first two months of life. Initially unable to fully regulate its own heart rate, immature infant heart rate is characterized by a steady decline in heart rate when separated from its parents, which is likely representative of a lack of thermoregulatory capacity. For approximately the first three months of age, infants are usually carried by their parents and unrelated helpers, presumably to maintain body temperature (Borjon & Ghazanfar, 2014; Harshaw, Blumberg, & Alberts, 2016). Later in development, after the first few months of life, infant heart rate stabilizes and does not exhibit the decline seen earlier. We interpret this stabilization of heart rate as an indicator of thermoregulatory capacity, however we did not measure temperature in these infants. Parallel to the stabilization of heart rate over time, there is an increase in average heart rate across development. While human infants, and indeed many mammals, exhibit a general deceleration in heart during maturation, it is common for very small mammals to experience a general increase in average heart rate from infancy to adulthood. In addition, the variability of marmoset infants' heart rate decreased over development, which is consistent with data from human infants (Richards & Cameron, 1989). Developmental changes in the spectral components of cardiac activity have generally been couched in terms of parasympathetic-sympathetic-vagal balance (Friedman, 2007; Pfaff, 2006; Porges & Lewis, 2009; Sayin, Chapuis, Chevalier, Barrès, & Julien, 2016). The notion that the branches of the autonomic nervous system are antagonistically related is quite old and likely outdated (Pfaff, 2006). Such an argument originated from original description of the iris in the human eye: sympathetic nerves cause it to dilate, while parasympathetic nerves cause it to constrict (Löwenstein & Loewenfeld, 1950). Other examples of balance between the two branches of the autonomic systems are rare (Pfaff, 2006).

In our study, we could quantify the maturation of the marmoset's arousal states by comparing various oscillations inherent to the autonomic nervous system. Over the first two months of life, we observed an increase in the magnitude of low frequency oscillations in heart rate, representing the Mayer wave, and a decrease in high frequency oscillations which are modulated by respiratory activity. Thus, we used the ratio of low to high frequency (LH ratio) as an index of spectral change in heart rate, the development of which is itself oscillatory. We correlated the development of this ratio to changes in motor and vocal behavior and demonstrate that, on days where the LH ratio of cardiac activity is high, infant marmosets spend more time exhibiting mature motor activity as well as lower entropy and thus more mature vocalizations. This demonstrates that developmental changes in arousal can be associated with corresponding changes in motor and vocal behavior during development.

To determine whether the relationships between arousal, motor activity, and vocal behavior are demonstrably causal, we used a variant of structural equation modelling, termed *pairwise LiNGAM* (Hyvärinen & Smith, 2013). Specifically designed to calculate causality without the presence of known priors, this algorithm leverages pairwise likelihood ratios to infer the causal direction between pairs of variables and provides the strength of the connection. Statistically, the LH ratio drives the development of mature motor activity which then drives vocal development. In addition, there is a connection between the LH ratio and vocal development. We argue for a sequential development of arousal dynamics, followed by motor activity, followed by vocal development, in which arousal is an important engine to facilitate the development of mature behaviors.

In this dissertation, I have argued that dynamic patterns of autonomic activity support impact animal behavior (Borjon et al., 2016; Borjon & Ghazanfar, 2014; Pfaff, 2006) and are important for primate vocal production across all stages of development (Ghazanfar et al., 2017). First, we showed that the timing and probability of adult marmoset vocal behavior is linked to the

slow, rhythmic fluctuation of the autonomic nervous system (Borjon et al., 2016). Second, we showed that during development, similar fluctuations in arousal are correlated to the maturation of motor behavior and vocal production. There is a statistically causal relationship between arousal development and both vocal and motor development. The extent to which fluctuations in arousal drive natural human behavior is unknown and would be a viable path of future study. As discussed in Chapter 1, many uniquely human behaviors are argued to be the result of complex cortical interactions and cognitive mechanisms (Borjon & Ghazanfar, 2014). There is growing evidence that non-cortical structures such as the limbic system (Chanes & Barrett, 2016) are fundamentally important cortical processing to occur and play an important role in human behavior. The results of this dissertation suggest that the often overlooked autonomic nervous system may critically contribute to the production and development of natural primate behaviors from infancy to adulthood.

Chapter 6

References

Adolph, E. F. (1971). Ontogeny of heart-rate controls in hamster, rat, and guinea pig. *The American Journal of Physiology*, *220*, 1896–1902.

An, X., Bandler, R., Ongür, D., & Price, J. L. (1998). Prefrontal cortical projections to longitudinal columns in the midbrain periaqueductal gray in Macaque monkeys. *Journal of Comparative Neurology*, *401*, 455–479.

Bastian, A., & Schmidt, S. (2008). Affect cues in vocalizations of the bat,Megaderma lyra,during agonistic interactions. *Journal of the Acoustical Society of America*, *124*, 598–608.

Bercovitch, F. B., Hauser, M. D., & Jones, J. H. (1995). The endocrine stress response and alarm vocalizations in rhesus macaques. *Animal Behaviour*, *49*, 1703–1706.

Berna, F., Goldberg, P., Horwitz, L. K., Brink, J., Holt, S., Bamford, M., & Chazan, M. (2012). Microstratigraphic evidence of in situ fire in the Acheulean strata of Wonderwerk Cave, Northern Cape province, South Africa. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, E1215–20.

Bernardi, L., Porta, C., Gabutti, A., Spicuzza, L., & Sleight, P. (2001). Modulatory effects of respiration. *Autonomic Neuroscience : Basic & Clinical*, *90*, 47–56.

Berntson, G. G., & Cacioppo, J. T. (2007). Integrative physiology: Homeostasis, allostasis and the orchestration of systemic physiology. *Handbook of Psychophysiology*, *3*, 433–452.

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). Respiratory sinus arrhythmia: autonomic origins, physiological mechanisms, and psychophysiological implications. *Psychophysiology*, *30*, 183–196.

Bezerra, B. M., & Souto, A. (2008). Structure and Usage of the Vocal Repertoire of Callithrix jacchus. *International Journal of Primatology*, *29*, 671–701.

Bhadoria, R., Watson, D., Danson, P., Ferrier, I. N., Mcallister, V. I., & Moore, P. B. (2003). Enlargement of the Third Ventricle in affective disorders. *Indian Journal of Psychiatry*, *45*, 147–150.

Blumstein, D. T., & Armitage, K. B. (1997). Alarm calling in yellow-bellied marmots: I. The meaning of situationally variable alarm calls. *Animal Behaviour*, *53*, 143–171.

Bonyhay, I., Risk, M., & Freeman, R. (2013). High-Pass Filter Characteristics of the Baroreflex – A Comparison of Frequency Domain and Pharmacological Methods. *PLoS One*, *8*, e79513–12.

Borjon, J. I., & Ghazanfar, A. A. (2014). Convergent evolution of vocal cooperation without convergent evolution of brain size. *Brain, Behavior and Evolution*, *84*, 93–102.

Borjon, J. I., Takahashi, D. Y., Cervantes, D. C., & Ghazanfar, A. A. (2016). Arousal dynamics drive vocal production in marmoset monkeys. *J Neurophysiol*, *116*, 753–764.

Boysen, S. T., & Berntson, G. G. (1989). Conspecific recognition in the chimpanzee (Pan troglodytes): cardiac responses to significant others. *Journal of Comparative Psychology (Washington, D.C. : 1983)*, *103*, 215–220.

Braun, K., Schultz-Darken, N., Schneider, M., Moore, C. F., & Emborg, M. E. (2015). Development of a novel postnatal neurobehavioral scale for evaluation of common marmoset monkeys. *American Journal of Primatology*, *77*, 401–417.

Bridges, R. S., Zarrow, M. X., Goldman, B. D., & Denenberg, V. H. (1974). A developmental study of maternal responsiveness in the rat. *Physiol Behav*, *12*, 149–151.

Briefer, E. F., Maigrot, A. L., Mandel, R., Freymond, S. B., Bachmann, I., & Hillmann, E. (2015).

Segregation of information about emotional arousal and valence in horse whinnies. *Scientific Reports*, *4*, 9989.
Burkart, J. M., & van Schaik, C. P. (2010). Cognitive consequences of cooperative breeding in primates? *Anim Cogn*, *13*, 1–19.
Burkart, J. M., Hrdy, S. B., & van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *18*, 175–186.
Byrge, L., Sporns, O., & Smith, L. B. (2014). Developmental process emerges from extended brain-body-behavior networks. *Trends in Cognitive Sciences*, *18*, 395–403.
Cacioppo, J. T., Uchino, B. N., & Berntson, G. G. (1994). Individual differences in the autonomic origins of heart rate reactivity: the psychometrics of respiratory sinus arrhythmia and preejection period. *Psychophysiology*, *31*, 412–419.
Cardinal, R. N., Parkinson, J. A., Hall, J., & Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience and Biobehavioral Reviews*, *26*, 321–352.
Cashon, C. H., Ha, O. R., Allen, C. L., & Barna, A. C. (2013). A U-Shaped Relation Between Sitting Ability and Upright Face Processing in Infants. *Child Development*, *84*, 802–809.
Champagne, F. A., & Curley, J. P. (2008). Maternal regulation of estrogen receptor α methylation. *Current Opinion in Pharmacology*, *8*, 735–739.
Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech., *5*, e1000436.
Chanes, L., & Barrett, L. F. (2016). Redefining the Role of Limbic Areas in Cortical Processing. *Trends in Cognitive Sciences*, *20*, 96–106.
Chapple, E. D. (1970). Culture and biological man.
Chen, X., Li, S., & Kirouac, G. J. (2014). Blocking of corticotrophin releasing factor receptor-1 during footshock attenuates context fear but not the upregulation of prepro-orexin mRNA in rats. *Pharmacology Biochemistry and Behavior*, *120*, 1–6.
Cheney, D. L., & Seyfarth, R. M. (1997). Reconciliatory grunts by dominant female baboons influence victims' behaviour. *Animal Behaviour*, *54*, 409–418.
Chiel, H. J., & Beer, R. D. (1997). The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, *20*, 553–557.
Choi, J. Y., Takahashi, D. Y., & Ghazanfar, A. A. (2015). Cooperative vocal control in marmoset monkeys via vocal feedback. *J Neurophysiol*, *114*, 274–283.
Clara, E., Tommasi, L., & Rogers, L. J. (2008). Social mobbing calls in common marmosets (Callithrix jacchus): effects of experience and associated cortisol levels. *Anim Cogn*, *11*, 349–358.
Coudé, G., Ferrari, P. F., Rodà, F., Maranesi, M., Borelli, E., Veroni, V., et al. (2011). Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. *PLoS One*, *6*, e26822–e26822.
Critchley, H. D., & Harrison, N. A. (2013). Visceral influences on brain and behavior. *Neuron*, *77*, 624–638.
Cross, N., & Rogers, L. J. (2004). Diurnal cycle in salivary cortisol levels in common marmosets. *Dev Psychobiol*, *45*, 134–139.
Cygankiewicz, I., & Zareba, W. (2013). Heart rate variability. *Handbook of Clinical Neurology*, *117*, 379–393.
Dettling, A. C., Feldon, J., & Pryce, C. R. (2002). Repeated parental deprivation in the infant common marmoset (Callithrix jacchus, primates) and analysis of its effects on early development. *Biological Psychiatry*, *52*, 1037–1046.
Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate cortex to

behaviour. *Brain*, *118*, 279–306.
Dunbar, R. (1998). Grooming, Gossip, and the Evolution of Language. Harvard University Press.
Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution., *36*, 562–572.
Ebitz, R. B., & Platt, M. L. (2015). Neuronal activity in primate dorsal anterior cingulate cortex signals task conflict and predicts adjustments in pupil-linked arousal. *Neuron*, *85*, 628–640.
Ebitz, R. B., Pearson, J. M., & Platt, M. L. (2014). Pupil size and social vigilance in rhesus macaques., *8*, 100.
Egnor, S. E. R., & Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neurosciences*, *27*, 649–654.
Elias, G., & Broerse, J. (1996). Developmental changes in the incidence and likelihood of simultaneous talk during the first two years: a question of function. *J. Child Lang.*, *23*, 201–217.
Elias, G., Hayes, A., & Broerse, J. (1986). Maternal control of co-vocalization and inter-speaker silences in mother-infant vocal engagement. *Journal of Child Psychology and Psychiatry*, *27*, 409–415.
Faes, L., Porta, A., & Nollo, G. (2015). Algorithms for the inference of causality in dynamic processes: Application to cardiovascular and cerebrovascular variability (pp. 1789–1792). Presented at the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE.
Farkas, E., Jansen, A. S., & Loewy, A. D. (1997). Periaqueductal gray matter projection to vagal preganglionic neurons and the nucleus tractus solitarius. *Brain Res*, *764*, 257–261.
Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., & Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behav Dev*, *34*, 569–577.
Fischer, J., Hammerschmidt, K., Cheney, D. L., & Seyfarth, R. M. (2001). Acoustic features of female chacma baboon barks. *Ethology*, *107*, 33–54.
Fitch, W. T., & Hauser, M. D. (1995). Vocal Production in Nonhuman-Primates - Acoustics, Physiology, and Functional Constraints on Honest Advertisement. *American Journal of Primatology*, *37*, 191–219.
Flack, J. C. (2013). Animal communication: hidden complexity. *Curr Biol*, *23*, R967–9.
Fragaszy, D., Visalberghi, E., & Galloway, A. (1997). Infant tufted capuchin monkeys' behaviour with novel foods: opportunism, not selectivity. *Animal Behaviour*, *53*, 1337–1343.
Franklin, B., Warlaumont, A. S., Messinger, D., Bene, E., Nathani Iyer, S., Lee, C.-C., et al. (2013). Effects of Parental Interaction on Infant Vocalization Rate, Variability and Vocal Type. *Language Learning and Development*, *10*, 279–296.
Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*, 1325–1330.
Friedman, B. H. (2007). An autonomic flexibility–neurovisceral integration model of anxiety and cardiac vagal tone. *Biological Psychology*, *74*, 185–199.
Fry, W. F., & Savin, W. M. (1988). Mirthful laughter and blood pressure. *Humor*, *1*, 49–62.
Furlong, T., & Carrive, P. (2007). Neurotoxic lesions centered on the perifornical hypothalamus abolish the cardiovascular and behavioral responses of conditioned fear to context but not of restraint. *Brain Res*, *1128*, 107–119.
Gammaitoni, L., Hänggi, P., Jung, P., & Marchesoni, F. (1998). Stochastic resonance. *Reviews of Modern Physics*, *70*, 223–287.
Garey, J., Goodwillie, A., Frohlich, J., Morgan, M., Gustafsson, J.-A., Smithies, O., et al. (2003). Genetic contributions to generalized arousal of brain and behavior. *Proc. Natl. Acad. Sci.*, *100*, 11019–11022.

Ghazanfar, A. A., & Rendall, D. (2008). Evolution of human vocal production. *Curr Biol*, *18*, R457–60.
Ghazanfar, A. A., & Takahashi, D. Y. (2014a). Facial expressions and the evolution of the speech rhythm. *Journal of Cognitive Neuroscience*, *26*, 1196–1207.
Ghazanfar, A. A., & Takahashi, D. Y. (2014b). The evolution of speech: vision, rhythm, cooperation. *Trends in Cognitive Science*, *18*, 543–553.
Ghazanfar, A. A., & Zhang, Y. S. (2016). The autonomic nervous system is the engine for vocal development through social feedback. *Curr Opin Neurobiol*, *40*, 155–160.
Ghazanfar, A. A., Smith-Rohrberg, D., Pollen, A. A., & Hauser, M. D. (2002). Temporal cues in the antiphonal long-calling behaviour of cottontop tamarins. *Animal Behaviour*, *64*, 427–438.
Ghazanfar, A. A., Takahashi, D. Y., Zhang, Y. S., & Borjon, J. I. (2017). Marmoset monkey vocal communication: Common trajectorieswith humans and possible mechanisms. In J. Ellison & M. Sera, *The Minnesota Symposium on Child Psychology* (Vol. 39). Wiley Publishing.
Ghazanfar, A., Flombaum, J., Miller, C., & Hauser, M. (2001). The units of perception in the antiphonal calling behavior of cotton-top tamarins (Saguinus oedipus): playback experiments with long calls. *Journal of Comparative Physiology. a, Neuroethology, Sensory, Neural, and Behavioral Physiology*, *187*, 27–35.
Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, *15*, 511–517.
Goldizen, A. W. (1990). A comparative perspective on the evolution of tamarin and marmoset social systems, *11*, 63–83.
Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol Sci*, *19*, 515–523.
Goldstein, M. H., Bornstein, M. H., Schwade, J. A., Baldwin, F., & Brandstadter, R. (2009). The value of vocalizing: five-month-old infants associate their own non-cry vocalizations with responses from caregivers. *Child Development*, *80*, 636–644.
Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences*, *100*, 8030–8035.
Gothard, K. M. (2014). The amygdalo-motor pathways and the control of facial expressions. *Frontiers in Neuroscience*, *8*, 42.
Gottlieb, G. (1992). Individual development & evolution: the genesis of novel behavior. New York: Oxford University Press.
Gracanin, A., Bylsma, L. M., & Vingerhoets, J. J. M. (2014). Is crying a self-soothing behavior? *Frontiers in Psychology*, *5*, 502.
Grafe, T. U. (1996). The function of call alternation in the African reed frog (Hyperolius marmoratus): precise call timing prevents auditory masking. *Behavioral Ecology and Sociobiology*.
Gray, M. A., Beacher, F. D., Minati, L., Nagai, Y., Kemp, A. H., Harrison, N. A., & Critchley, H. D. (2012). Emotional appraisal is influenced by cardiac afferent information. *Emotion*, *12*, 180–191.
Gray, P., & Brooks, P. J. (1984). Effect of lesion location within the medial preoptic-anterior hypothalamic continuum on maternal and male sexual behaviors in female rats. *Behav Neurosci*, *98*, 703–711.
Greenfield, M. D. (1994). Synchronous and Alternating Choruses in Insects and Anurans: Common Mechanisms and Diverse Functions. *American Zoologist*, *34*, 605–615.
Gros-Louis, J., West, M. J., & King, A. P. (2010). Comparative perspectives on the missing link: Communicative pragmatics. In M. S. Blumberg, J. H. Freeman, & S. R. Robinson, *The Oxford handbook of developmental behavioral neuroscience* (pp. 684–707). Oxford, UK: Oxford University Press.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, *30*, 509–516.

Grossman, P., & Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*, *74*, 263–285.

Hage, S. R., & Nieder, A. (2013). Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nature Communications*, *4*, 2409–2409.

Hage, S. R., Gavrilov, N., & Nieder, A. (2013). Cognitive control of distinct vocalizations in rhesus monkeys. *Journal of Cognitive Neuroscience*, *25*, 1692–1701.

Hammerschmidt, K., & Fischer, J. (2008). Constraints in primate vocal production. In G. U & O. K, (pp. 93–119). Cambridge, MA: MIT Press.

Hare, B., Melis, A. P., Woods, V., Hastings, S., & Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Current Biology*, *17*, 619–623.

Harshaw, C., & Lickliter, R. (2011). Biased embryos: Prenatal experience alters the postnatal malleability of auditory preferences in bobwhite quail. *Dev Psychobiol*, *53*, 291–302.

Harshaw, C., Blumberg, M. S., & Alberts, J. R. (2016). Thermoregulation, energetics, and behavior. (pp. 931–952). Washington: American Psychological Association.

He, M., Walle, E. A., & Campos, J. J. (2015). A Cross-National Investigation of the Relationship Between Infant Walking and Language Development. *Infancy*, *20*, 283–305.

Henry, B. L., Minassian, A., Paulus, M. P., Geyer, M. A., & Perry, W. (2010). Heart rate variability in bipolar mania and schizophrenia. *Journal of Psychiatric Research*, *44*, 168–176.

Herculano-Houzel, S., Collins, C. E., Wong, P., & Kaas, J. H. (2007). Cellular scaling rules for primate brains. *Proceedings of the National Academy of Sciences*, *104*, 3562–3567.

Hewes, G. W. (1992). Primate Communication and the Gestural Origin of Language. *Current Anthropology*, *33*, 65–84.

Hilbrink, E. E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: a longitudinal study of mother-infant interaction. *Frontiers in Psychology*, *6*, 1492.

Hoffman, C. M., Flory, G. S., & Alberts, J. R. (1999). Ontogenetic adaptation and learning: A developmental constraint in learning for a thermal reinforcer. *Dev Psychobiol*, *34*, 73–86.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*. doi:10.2307/4615733

Hrdy, S. B. (2005). Evolutionary context of human development: The cooperative breeding model. In C. S. Carter, L. Ahnert, K. E. Grossmann, S. B. Hrdy, M. E. Lamb, S. W. Porges, & N. Sachser, *Attachment and Bonding: A New Synthesis* (pp. 39–68). Cambridge, MA: MIT Press.

Hsu, H.-C., & Fogel, A. (2001). Infant Vocal Development in a Dynamic Mother-Infant Communication System. *Infancy*, *2*, 87–109.

Hyvärinen, A., & Smith, S. M. (2013). Pairwise likelihood ratios for estimation of non-Gaussian structural equation models. *The Journal of Machine Learning Research*, *14*, 111–152.

Iverson, J. M. (2010). Developing language in a developing body: the relationship between motor development and language development. *Journal of Child Language*, *37*, 229–261.

Iverson, J. M., & Braddock, B. A. (2011). Gesture and Motor Skill in Relation to Language in Children With Language Impairment. *Journal of Speech Language and Hearing Research*, *54*, 72–86.

Iverson, J. M., & Fagan, M. K. (2004). Infant Vocal–Motor Coordination: Precursor to the Gesture–Speech System? *Child Development*, *75*, 1053–1066.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P., & Stern, D. N. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for*

Research in Child Development, *66*, i–149.
Jasnow, M., & Feldstein, S. (1986). Adult-like temporal characteristics of mother-infant vocal interactions., *57*, 754–761.
Jennings, J. R., van der Molen, M. W., Somsen, R. J., & Terezis, C. (1990). On the shift from anticipatory heart rate deceleration to acceleratory recovery: revisiting the role of response factors. *Psychophysiology*, *27*, 385–395.
Jerison, H. (2012). Evolution of the Brain and Intelligence.
Julien, C. (2006). The enigma of Mayer waves: Facts and models. *Cardiovascular Research*, *70*, 12–21.
Jürgens, U. (2009). The neural control of vocalization in mammals: a review. *Journal of Voice : Official Journal of the Voice Foundation*, *23*, 1–10.
Kajikawa, S., Amano, S., & Kondo, T. (2004). Speech overlap in Japanese mother-child conversations. *J. Child Lang.*, *31*, 215–230.
Kale, D. C., Che, Z., Bahadori, M. T., Li, W., Liu, Y., & Wetzel, R. (2015). Causal Phenotype Discovery via Deep Networks. *AMIA Annual Symposium Proceedings*, *2015*, 677–686.
Kampe, J., Tschöp, M. H., Hollis, J. H., & Oldfield, B. J. (2009). An anatomic basis for the communication of hypothalamic, cortical and mesolimbic circuitry in the regulation of energy balance. *European Journal of Neuroscience*, *30*, 415–430.
Katz, P. S. (2011). Neural mechanisms underlying the evolvability of behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 2086–2099.
Kayaba, Y., Nakamura, A., Kasuya, Y., Ohuchi, T., Yanagisawa, M., Komuro, I., et al. (2003). Attenuated defense response and low basal blood pressure in orexin knockout mice. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, *285*, R581–R593.
Kent, R. D., & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *The Journal of the Acoustical Society of America*, *72*, 353–365.
Kitchen, D. M., Cheney, D. L., & Seyfarth, R. M. (2005). Male chacma baboons (Papio cynocephalus ursinus) discriminate loud call contests between rivals of different relative ranks. *Anim Cogn*, *8*, 1–6.
Kleinow, J., & Smith, A. (2006). Potential interactions among linguistic, autonomic, and motor factors in speech. *Dev Psychobiol*, *48*, 275–287.
Kondo, Izawa, Watanabe. (2010). Perceptual mechanism for vocal individual recognition in jungle crows (Corvus macrorhynchos): contact call signature and discrimination. *Behaviour*, *147*, 1051–1072.
Krubitzer, L., & Kaas, J. (2005). The evolution of the neocortex in mammals: how is phenotypic diversity generated? *Curr Opin Neurobiol*, *15*, 444–453.
Kulahci, I. G., Rubenstein, D. I., & Ghazanfar, A. A. (2015). Lemurs groom-at-a-distance through vocal networks. *Animal Behaviour*, *In revision*.
Lacey, J. I. (1967). Somatic response patterning and stress: Some revisions of activation theory. Psychological stress: Issues in research.
Latalova, K., Prasko, J., Diveky, T., Grambal, A., Kamaradova, D., Velartova, H., et al. (2010). Autonomic nervous system in euthymic patients with bipolar affective disorder. *Neuro Endocrinology Letters*, *31*, 829–836.
Law, A. J., Pei, Q., Feldon, J., Pryce, C. R., & Harrison, P. J. (2009). Gene expression in the anterior cingulate cortex and amygdala of adolescent marmoset monkeys following parental separations in infancy. *International Journal of Neuropsychopharmacology*, *12*, 761–772.
Lecanuet, J.-P., Fifer, W. P., Krasnegor, N. A., & Smotherman, W. P. (2013). Fetal Development. Psychology Press.
Levy, B. (2013). Autonomic nervous system arousal and cognitive functioning in bipolar disorder. *Bipolar Disorders*, *15*, 70–79.

Lewin, W., & Whitty, C. W. M. (1960). Effects of anterior cingulate stimulation in conscious human subjects. *J Neurophysiol*, *23*, 445–447.
Libertus, K., & Needham, A. (2011). Reaching experience increases face preference in 3-month-old infants. *Developmental Science*, *14*, 1355–1364.
Libertus, K., & Violi, D. A. (2016). Sit to Talk: Relation between Motor Skills and Language Development in Infancy. *Frontiers in Psychology*, *7*, 527–8.
Linden, W. (1987). A microanalysis of autonomic activity during human speech. *Psychosomatic Medicine*, *49*, 562–578.
Loewy, A. D., & Spyer, K. M. (1990). Central regulation of autonomic functions. Oxford University Press New York.
Logue, D. M., Chalmers, C., & Gowland, A. H. (2008). The behavioural mechanisms underlying temporal coordination in black-bellied wren duets. *Animal Behaviour*, *75*, 1803–1808.
Löwenstein, O., & Loewenfeld, I. E. (1950). Mutual role of sympathetic and parasympathetic in shaping of the pupillary reflex to light. Arch. Neurol. Psychiat.(Chic.).
Lynch, J. J., Thomas, S. A., & Long, J. M. (1980). Human speech and blood pressure. *The Journal of* …. doi:10.1177/00238309060490030201
M, M., JR, W., H, R., W, H., HJ, M., A, T., et al. (1992). Developmental biology of the common marmoset: proposal for a "postnatal staging". *Journal of Medical Primatology*, *21*, 285–298.
MacNeilage, P. F., & Davis, B. L. (2000). On the origin of internal structure of word forms. *Science*, *288*, 527–531.
MacNeilage, P. F., Davis, B. L., & Kinney, A. (1999). Origin of serial-output complexity in speech. *Psychological* ….
Malinow, K. L., Lynch, J. J., Foreman, P. J., Friedmann, E., & Thomas, S. A. (1986). Blood pressure increases while signing in a deaf population. *Psychosomatic Medicine*, *48*, 95–101.
Manser, M. B., Bell, M. B., & Fletcher, L. B. (2001). The information that receivers extract from alarm calls in suricates. *Proceedings. Biological Sciences / the Royal Society*, *268*, 2485–2491.
Masataka, N., & Biben, M. (1987). Temporal Rules Regulating Affiliative Vocal Exchanges of Squirrel Monkeys. *Behaviour*, *101*, 311–319.
McEwen, B. S., & Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior*, *43*, 2–15.
McMeekin, H. (2002). Autonomic peripheral vascular dysregulation and mood disorder. Journal of affective disorders.
McRoberts, G. W., Studdert-Kennedy, M., & Shankweiler, D. P. (1995). The role of fundamental frequency in signaling linguistic stress and affect: Evidence for a dissociation. *Perception & Psychophysics*, *57*, 159–174.
Miller, C. T., & Thomas, A. W. (2012). Individual recognition during bouts of antiphonal calling in common marmosets. *Journal of Comparative Physiology. a, Neuroethology, Sensory, Neural, and Behavioral Physiology*, *198*, 337–346.
Miller, C. T., & Wang, X. (2006). Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. *Journal of Comparative Physiology. a, Neuroethology, Sensory, Neural, and Behavioral Physiology*, *192*, 27–38.
Mitani, J. C. (1985). Gibbon Song Duets and Intergroup Spacing. *Behaviour*, *92*, 59–96.
Montano, N., Lombardi, F., Gnecchi Ruscone, T., Contini, M., Finocchiaro, M. L., Baselli, G., et al. (1992). Spectral analysis of sympathetic discharge, R-R interval and systolic arterial pressure in decerebrate cats. *Journal of the Autonomic Nervous System*, *40*, 21–31.
Moriceau, S., & Sullivan, R. M. (2005). Neurobiology of infant attachment. *Dev Psychobiol*, *47*, 230–242.

Morin, D., & Viala, D. (2002). Coordinations of locomotor and respiratory rhythms in vitro are critically dependent on hindlimb sensory inputs. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, *22*, 4756–4765.
Nakahara, F., & Miyazaki, N. (2011). Vocal exchanges of signature whistles in bottlenose dolphins (Tursiops truncatus). *Journal of Ethology*, *29*, 309–320.
Nattie, E., & Li, A. (2012). Respiration and autonomic regulation and orexin. In *Orexin/Hypocretin System* (Vol. 198, pp. 25–46). Elsevier.
Nieuwenhuys, R., Donkelaar, ten, H. J., & Nicholson, C. (2014). The Central Nervous System of Vertebrates. Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-18262-4
Norcross, J. L., & Newman, J. D. (1993). Context and gender-specific differences in the acoustic structure of common marmoset (Callithrix jacchus) phee calls. *American Journal of Primatology*, *30*, 37–54.
Norcross, J. L., & Newman, J. D. (1999). Effects of separation and novelty on distress vocalizations and cortisol in the common marmoset (Callithrix jacchus). *American Journal of Primatology*, *47*, 209–222.
Numan, M. (1994). A neural circuitry analysis of maternal behavior in the rat. *Acta Paediatrica*, *83*, 19–28.
O'Conaill, B., Whittaker, S., & Wilbur, S. (2009). Conversations Over Video Conferences: An Evaluation of the Spoken Aspects of Video-Mediated Communication. *Human-Computer Interaction*, *8*, 389–428.
Obrist, P. A. (1981). Cardiovascular psychophysiology: A perspective.
Obrist, P. A., Light, K. C., McCubbin, J. A., Hutcheson, J. S., & Hoffer, J. L. (1979). Pulse transit time: relationship to blood pressure and myocardial performance. *Psychophysiology*, *16*, 292–301.
Obrist, P. A., Webb, R. A., & Sutterer, J. R. (1969). Heart rate and somatic changes during aversive conditioning and a simple reaction time task. *Psychophysiology*, *5*, 696–723.
Obrist, P. A., Webb, R. A., Sutterer, J. R., & Howard, J. L. (1970). The Cardiac-Somatic Relationship: Some Reformulations. *Psychophysiology*, *6*, 569–587.
Oller, D. K. (2000). The emergence of the speech capacity. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Oullier, O., de Guzman, G. C., Jantzen, K. J., Lagarde, J., & Kelso, J. A. S. (2008). Social coordination dynamics: Measuring human bonding. *Social Neuroscience*, *3*, 178–192.
Owings, D. H., & Virginia, R. A. (1978). Alarm calls of California ground squirrels (Spermophilus beecheyi). *Zeitschrift Für Tierpsychologie*, *46*, 58–70.
Owren, M. J., Amoss, R. T., & Rendall, D. (2011). Two organizing principles of vocal production: Implications for nonhuman and human primates. *American Journal of Primatology*, *73*, 530–544.
Owren, M. J., Rendall, D., & Ryan, M. J. (2010). Redefining animal signaling: influence versus information in communication. *Biology & Philosophy*, *25*, 755–780.
Padberg, J., Disbrow, E., & Krubitzer, L. (2005). The organization and connections of anterior and posterior parietal cortex in titi monkeys: do New World monkeys have an area 2? *Cereb Cortex*, *15*, 1938–1963.
Padberg, J., Franca, J. G., Cooke, D. F., Soares, J. G. M., Rosa, M. G. P., Fiorani, M., et al. (2007). Parallel evolution of cortical areas involved in skilled hand use. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, *27*, 10106–10115.
Paus, T. S. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nat Rev Neurosci*, *2*, 417–424.
Pellis, S. M., & Bell, H. C. (2011). Closing the circle between perceptions and behavior: A cybernetic view of behavior and its consequences for studying motivation and development. *Developmental*

Cognitive Neuroscience, *1*, 404–413.
Pereira, M., Andreatini, R., Schwarting, R. K. W., & Brenes, J. C. (2014). Amphetamine-induced appetitive 50-kHz calls in rats: a marker of affect in mania?, *231*, 2567–2577.
Peters, H. F. M., & Hulstijn, W. (1984). Stuttering and anxiety: The difference between stutterers and nonstutterers in verbal apprehension and physiologic arousal during the anticipation of speech and non-speech tasks. *Journal of Fluency Disorders*, *9*, 67–84.
Pfaff, D. (2006). Brain arousal and information theory. Cambridge, MA: Harvard University Press.
Phillips, I. R. (1976). Skeletal development in the foetal and neonatal marmoset (Callithrix jacchus). *Laboratory Animals*, *10*, 317–333.
Porges, S. W., & Lewis, G. F. (2009). The polyvagal hypothesis: common mechanisms mediating autonomic regulation, vocalizations and listening. In B. SM, *Handbook of Behavioral Neuroscience* (Vol. 19, pp. 255–264). London: Academic Press.
Potts, J. T. (2002). Neural circuits controlling cardiorespiratory responses: baroreceptor and somatic afferents in the nucleus tractus solitarius. *Clinical and Experimental Pharmacology & Physiology*, *29*, 103–111.
Pryce, C. R., Aubert, Y., Maier, C., Pearce, P. C., & Fuchs, E. (2011). The developmental impact of prenatal stress, prenatal dexamethasone and postnatal social stress on physiology, behaviour and neuroanatomy of primate offspring: studies in rhesus macaque and common marmoset. *Psychopharmacology*, *214*, 33–53.
Remedios, R., Logothetis, N. K., & Kayser, C. (2009). An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. *Journal of Neuroscience*, *29*, 1034–1045.
Rendall, D. (2003). Acoustic correlates fo caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *Journal of the Acoustical Society of America*, *113*, 3390–3402.
Rendall, D., & Owren, M. J. (2009). Vocalizations as tools for influencing the affect and behavior of others. In S. M. Brudzynski, *Handbook of mammalian vocalization: an integrative neuroscience approach* (Vol. 19, pp. 177–185). Academic Press.
Rendall, D., Owren, M. J., & Ryan, M. J. (2009). What do animal signals mean? *Animal Behaviour*, *78*, 233–240.
Reyes del Paso, G. A., Langewitz, W., Mulder, L. J. M., van Roon, A., & Duschek, S. (2013). The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: A review with emphasis on a reanalysis of previous studies. *Psychophysiology*, *50*, 477–487.
Richards, J. E., & Cameron, D. (1989). Infant heart-rate variability and behavioral developmental status. *Infant Behav Dev*, *12*, 45–58.
Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, *21*, 188–194.
Rosenblatt, J. S., Hazelwood, S., & Poole, J. (1996). Maternal Behavior in Male Rats: Effects of Medial Preoptic Area Lesions and Presence of Maternal Aggression. *Hormones and Behavior*, *30*, 201–215.
Rukstalis, M., Fite, J. E., & French, J. A. (2003). Social Change Affects Vocal Structure in a Callitrichid Primate (Callithrix kuhlii). *Ethology*, *109*, 327–340.
Ryan, M. J. (1988). Energy, Calling, and Selection. *Integrative and Comparative Biology*, *28*, 885–898.
Samson, W. K., Gosnell, B., Chang, J.-K., Resch, Z. T., & Murphy, T. C. (1999). Cardiovascular regulatory actions of the hypocretins in brain. *Brain Res*, *831*, 248–253.
Santos, C. V., French, J. A., & Otta, E. (1997). Infant Carrying Behavior in Callitrichid Primates: Callithrix and Leontopithecus. *International Journal of Primatology*, *18*, 889–907.
Sayin, H., Chapuis, B., Chevalier, P., Barrès, C., & Julien, C. (2016). Assessment of cardiac autonomic tone in conscious rats. *Autonomic Neuroscience : Basic & Clinical*, *194*, 26–31.
Schehka, S., Esser, K.-H., & Zimmermann, E. (2007). Acoustical expression of arousal in conflict

situations in tree shrews (Tupaia belangeri). *Journal of Comparative Physiology. a, Neuroethology, Sensory, Neural, and Behavioral Physiology*, *193*, 845–852.
Scheumann, M., Roser, A. E., Konerding, W., Bliech, E., Hedrich, H. J., & Zimmermann, E. (2012). Vocal correlates of sender-identity and arousal in the isolation calls of domestic kitten (Felis sylvestris catus). *Frontiers in Zoology*, *9*, 36.
Schmidt, R., & Morr, S. (2010). Coordination dynamics of natural social interactions. *International Journal of Sport Psychology*.
Schnell, C. R., & Wood, J. M. (1993). Measurement of blood pressure and heart rate by telemetry in conscious, unrestrained marmosets. *The American Journal of Physiology*, *264*, H1509–16.
Schultz-Darken, N., Braun, K. M., & Emborg, M. E. (2015). Neurobehavioral development of common marmoset monkeys. *Dev Psychobiol*, *58*, 141–158.
Sellen, A. J. (1995). Remote Conversations: The Effects of Mediating Talk With Technology. *HumanComputer Interaction*, *10*, 401–444.
Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology*, *54*, 145–173.
Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, *28*, 1070–1094.
Shahack-Gross, R., Berna, F., Karkanas, P., Lemorini, C., Gopher, A., & Barkai, R. (2014). Evidence for the repeated use of a central hearth at Middle Pleistocene (300 ky ago) Qesem Cave, Israel. *Journal of Archaeological Science*, *44*, 12–21.
Shahid, I. Z., Rahman, A. A., & Pilowsky, P. M. (2012). Orexin and Central Regulation of Cardiorespiratory System. *Vitamins and Hormones*, *89*, 159.
Shimazaki, H., & Shinomoto, S. (2010). Kernel bandwidth optimization in spike rate estimation. *Journal of Computational Neuroscience*, *29*, 171–182.
Smith, E. E., Guyton, A. C., Manning, R. D., & White, R. J. (1976). Integrated mechanisms of cardiovascular response and control during exercise in the normal human. *Progress in Cardiovascular Diseases*, *18*, 421–444.
Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, *7*, 343–348.
Snowdon, C. T., & Cronin, K. A. (2007). Cooperative breeders do cooperate. *Behavioural Processes*, *76*, 138–141.
Somsen, R. J. M., Jennings, J. R., & van der Molen, M. W. (2004). The cardiac cycle time effect revisited: temporal dynamics of the central-vagal modulation of heart rate in human reaction time tasks. *Psychophysiology*, *41*, 941–953.
Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in development: Motor skill acquisition facilitates three-dimensional object completion. *Developmental Psychology*, *46*, 129–138.
Sörös, P., Sokoloff, L. G., Bose, A., McIntosh, A. R., Graham, S. J., & Stuss, D. T. (2006). Clustered functional MRI of overt speech production. *NeuroImage*, *32*, 376–387.
Stanaszek, W. F., Kellerman, D., Brogden, R. N., & Romankiewicz, J. A. (1983). Prazosin update. A review of its pharmacological properties and therapeutic use in hypertension and congestive heart failure. *Drugs*, *25*, 339–384.
Stauss, H. M., Persson, P. B., Johnson, A. K., & Kregel, K. C. (1997). Frequency-response characteristics of autonomic nervous system function in conscious rats. *The American Journal of Physiology*, *273*, H786–95.
Steiper, M. E., & Young, N. M. (2006). Primate molecular divergence dates., *41*, 384–394.
Stephan, H., Baron, G., & Schwerdtfeger, W. K. (2012). The brain of the common marmoset (Callithrix jacchus): a stereotaxic atlas.
Stern, D. N., Jaffe, J., Beebe, B., & Bennett, S. L. (1975). Vocalizing in unison and in alternation: two

modes of communication within the mother-infant dyad. *Annals of the New York Academy of Sciences*, *263*, 89–100.
Stewart, A. M., Lewis, G. F., Heilman, K. J., Davila, M. I., Coleman, D. D., Aylward, S. A., & Porges, S. W. (2013). The covariation of acoustic features of infant cries and autonomic state. *Physiol Behav*, *120*, 203–210.
Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, *106*, 10587–10592.
Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour*, *55*, 673–687.
Syal, S., & Finlay, B. L. (2011). Thinking outside the cortex: social motivation in the evolution and development of language. *Developmental Science*, *14*, 417–430.
Takahashi, D. Y., Fenley, A. R., & Ghazanfar, A. A. (2016). Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*. doi:10.1098/rstb.2015.0370
Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Holmes, P., Borjon, J. I., & Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science*, *349*, 734–738.
Takahashi, D. Y., Narayanan, D. Z., & Ghazanfar, A. A. (2013). Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. *Current Biology*, *23*, 2162–2168.
Tardif, S. D., Layne, D. G., Cancino, L., & Smucny, D. A. (2002). Neonatal behavioral scoring of common marmosets (Callithrix jacchus): relation to physical condition and survival. *Journal of Medical Primatology*, *31*, 147–151.
Tchernichovski, O., & Marcus, G. (2014). ScienceDirect Vocal learning beyond imitation: mechanisms of adaptive vocal development in songbirds and human infants. *Curr Opin Neurobiol*, *28*, 42–47.
Thelen, E. (1979). Rhythmical stereotypies in normal human infants. *Animal Behaviour*, *27*, 699–715.
Thelen, E. (1981a). Kicking, rocking, and waving: Contextual analysis of rhythmical stereotypies in normal human infants. *Animal Behaviour*, *29*, 3–11.
Thelen, E. (1981b). Rhythmical behavior in infancy: An ethological perspective. *Developmental Psychology*, *17*, 237–257.
Thelen, E. (1991). Motor aspects of emergent speech: A dynamic approach. In N. A. Krasnegor, D. M. Rumbaugh, & R. L. Schiefelbusch, *Biological and behavioral determinants of language development* (pp. 339–362). Hillsdale, NJ: Lawrence Erlabaum Associates.
Thelen, E., & Smith, L. B. (2007). Dynamic Systems Theories. Hoboken, NJ, USA: John Wiley & Sons, Inc. doi:10.1002/9780470147658.chpsy0106
Tomasello, M. (2008). Origins of human communication. Cambridge, MA: MIT Press.
Tursky, B., Schwartz, G. E., & Crider, A. (1970). Differential patterns of heart rate and skin resistance during a digit-transformation task. *Journal of Experimental Psychology*, *83*, 451–457.
Uno, H., Tarara, R., Else, J. G., & Suleman, M. A. (1989). Hippocampal damage associated with prolonged and fatal stress in primates. *Journal of* ….
van de Borne, P., Rahnama, M., Mezzetti, S., Montano, N., Porta, A., Degaute, J. P., & Somers, V. K. (2001). Contrasting effects of phentolamine and nitroprusside on neural and cardiovascular variability. *AJP: Heart and Circulatory Physiology*, *281*, H559–65.
Van Kammen, D. P., & Murphy, D. L. (1975). Attenuation of the euphoriant and activating effects of d- and l-amphetamine by lithium carbonate treatment. *Psychopharmacologia*, *44*, 215–224.
Van Puyvelde, M., Loots, G., Meys, J., Neyt, X., Mairesse, O., Simcock, D., & Pattyn, N. (2015). Whose clock makes yours tick? How maternal cardiorespiratory physiology influences newborns' heart rate variability. *Biological Psychology*, *108*, 132–141.

Vitale, A., & Queyras, A. (1997). The response to novel foods in common marmoset (Callithrix jacchus): the effects of different social contexts. *Ethology*. doi:10.1111/j.1439-0310.1997.tb00155.x

Voelkl, B. (2006). Social contact influences the response of infant marmosets towards novel food. *Animal Behaviour*, *72*, 365–372.

Walle, E. A., & Campos, J. J. (2014). Infant language development is related to the acquisition of walking. *Developmental Psychology*, *50*, 336–348.

Wang, Y., Fang, Q., & Gong, N. (2014). Motor assessment of developing common marmosets. *Neuroscience Bulletin*, *30*, 387–393.

Wass, S. V., de Barbaro, K., & Clackson, K. (2015). Tonic and phasic co-variation of peripheral arousal indices in infants. *Biological Psychology*, *111*, 26–39.

Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody—influence of acoustic parameters and arousal. *NeuroImage*, *39*, 885–893.

Willson, M. C., Bell, E. C., Dave, S., Asghar, S. J., McGrath, B. M., & Silverstone, P. H. (2005). Valproate attenuates dextroamphetamine-induced subjective changes more than lithium. *European Neuropsychopharmacology*, *15*, 633–639.

Yamamoto, M. E., & de Araújo Lopes, F. (2004). Effect of Removal from the Family Group on Feeding Behavior by Captive Callithrix jacchus. *International Journal of Primatology*, *25*, 489–500.

Yosida, S., Kobayasi, K. I., Ikebuchi, M., Ozaki, R., & Okanoya, K. (2007). Antiphonal Vocalization of a Subterranean Rodent, the Naked Mole-Rat (Heterocephalus glaber). *Ethology*, *113*, 703–710.

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*, 244–262.

Zahed, S. R., Prudom, S. L., Snowdon, C. T., & Ziegler, T. E. (2008). Male parenting and response to infant stimuli in the common marmoset (Callithrix jacchus). *American Journal of Primatology*, *70*, 84–92.

Zeskind, P. S. (2013). Infant crying and the synchrony of arousal. In E. Altenmüller, S. Schmidt, & E. Zimmermann, (pp. 155–174). Oxford: OUP Oxford.

Zhang, W., Fukuda, Y., & Kuwaki, T. (2005). Respiratory and cardiovascular actions of orexin-A in mice. *Neuroscience Letters*, *385*, 131–136.

Zhang, Y. S., & Ghazanfar, A. A. (2016). Perinatally Influenced Autonomic System Fluctuations Drive Infant Vocal Sequences. *Curr Biol*, *26*, 1249–1260.