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

2015

THE ORGANIZATION OF ACTION IN THE FETUS: Prenatal development of orofacial movements in marmoset monkeys

Narayanan, Darshana Z

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01qr46r313d

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

THE ORGANIZATION OF ACTION IN THE FETUS

PRENATAL DEVELOPMENT OF OROFACIAL MOVEMENTS IN MARMOSET MONKEYS

Darshana Z Narayanan

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

AS ADVISED BY: Asif Ghazanfar

JUNE 2015

© Copyright by Darshana Z Narayanan, 2015. All rights reserved.

Abstract

In primates, a number of vocal behaviors are present at birth. These behaviors are often dubbed as ‘innate’ but how they develop remains mysterious. As opposed to simply appearing de novo at birth, we hypothesize that vocalizations begin to self-organize prenatally through the interactions between spontaneous neural activity, fetal biomechanics and the physical constraints of the uterine environment. We investigated how fetal movements relate to the postnatal mouth movements required for vocal production by performing ultrasound imaging on awake, pregnant marmoset monkeys (*Callithrix jacchus*). Our expectations were two-fold (1) body parts linked to different behaviors become increasingly differentiated; and (2) body parts used for the same behavior become increasingly coordinated. We found that orofacial and head movements occur together early in gestation but gradually decouple and move independently by the end of gestation. Critically, we also observed that signature features of marmoset infant calls emerge prenatally as distinct patterns of orofacial movements: Late in gestation, the duration and syllable number of the phee call and twitter call — calls produced by marmoset monkeys on the first postnatal day — are evident in the orofacial movements of fetal marmosets. Our study shows that aspects of vocal behaviors in marmosets have a period of prenatal development.

To the Ghazanfar Lab

tighter than spandex on a fat kid

Acknowledgements

The study described in Chapter1 was done in collaboration with Asif A Ghazanfar, Daniel Y Takahashi, Lauren M Kelly and Sabina I Hlavaty. The study in Chapter2 was done in collaboration with Asif A Ghazanfar, Daniel Y Takahashi, Lauren M Kelly, Junzhou Huang, Department of Computer Science and Engineering, University of Texas at Arlington and Lin Yang, Department of Computer Science, University of Kentacky. I thank them all for their scientific contributions and for making this process an enjoyable one.

Table of Contents

General Introduction 1

Chapter 1: Self-organization of mouth and head modules in fetal marmosets

Summary 7

Introduction 7

Results 12

Discussion 30

Methods 38

Chapter 2: The prenatal origins of "innate" vocalizations in marmoset monkeys

Summary 45

Introduction 45

Results 48

Discussion 67

Methods 76

General Discussion 84

References 90

General Introduction

In western tradition, a baby is age zero the day it is born. This calculation is flawed. A baby is the result of 9 months of fetal development; the clock should start in-utero. Extensive studies on non-human species and work in recent years on human fetuses have established that life for a fetus is a rich and multifaceted experience. Paying attention to fetal development can give us a more complete understanding of the ontogeny of a behavior (Lehrman, 1953). For example, it is known that human neonates have the remarkable ability to distinguish their mother's face from the face an unknown female within the first few days of birth (Sai, 2005). Such rapid learning is facilitated by prenatal learning of the mother's voice — a manipulation study conducted in the hospitals of Malaysia demonstrates that babies who did not receive simultaneous exposure to the face and voice of the mother, failed to distinguish her face from that of a stranger (Sai, 2005). Prenatal learning of the mother's voice (DeCasper and Fifer, 1980) is only one of many examples of fetal learning.

Fetuses prefer their native language over a foreign language (Moon et al., 1993). Babies are significantly soothed and are more attentive to music their mother listened to during the last three months of pregnancy (Hepper, 1997). Neonates display an attraction to the odor of their amniotic fluid (Schaal and Rouby, 1990) and, prenatal flavor learning has been demonstrated (Smotherman and Robinson, 1985; Mennella et al., 2001. In-utero taste and odor aversion conditioning experiments show that, as young as embryonic day 17, rats can acquire an aversion that lasts until after weaning (Stickrod, 1982; Gruest et al., 2004). In animals, we have been able to tease apart the mechanisms underlying fetal learning. The importance of proprioceptive feedback in early motor development can be seen in the studies of Robinson and Smotherman (Smotherman and Robinson, 1988; Robinson and Smotherman, 1992). In their experiments, interlimb yokes were used to alter kinesthetic feedback from the limbs of rat fetuses. Fetal rats that were given the yoke training exhibited consistent changes in the coordination of their yoked limbs but not the other limbs. This result reveals that fetuses have a kinesthetic sense and can use proprioceptive feedback to modify motor behavior (Robinson and Kleven, 2005). These

are only a few examples demonstrating that prenatal experience is important for life after birth.

Further demonstrating the continuity between the fetus and its postnatal life, fetal behavior can predict postnatal behavior. Gross fetal movements can predict rates of crying after birth (St James-Roberts and Menon-Johansson, 1999) and more active fetuses become more active toddlers (Groome et al., 1999; DiPietro et al., 2001). The advent of four-dimensional ultrasound (4D-US) has opened up the possibility of observing human fetal activity in considerable detail and all behaviors exhibited by neonates have been observed at some point during the fetal period (DiPietro, 2008). Images produced by 4D-US scans show that general fetal movement is seen as early as the 8th week of pregnancy. During the second trimester of pregnancy (weeks 13-27) human fetuses display a wide range of movements that are initially chaotic but gradually become more organized. These movements include opening and closing of the mouth, swallowing, tongue protrusion, eye movements and movements of the whole body and/or limbs. By the 18th week the fetus regularly touches its face (e.g. eye rubbing, touching the mouth and ears) and facial expressions can be detected. Facial expressions (smiling, scowling, pouting, hiccupping, yawning) are very distinct and frequent by the end of gestation (de Vries et al., 1982, 1985; Prechtl, 1985; Kurjak et al., 2005; Berivoj Miskovic, 2011).

Of course, the prenatal environment is a restricted place. Behavior in such a confined environment does not always appear identical in the infant and fetus. Pecking behavior in chicks, for example, has been carefully studied and illustrates this point. Domestic chicks begin pecking at food grains etc. soon after hatching. The behavior consists of head-lunging, bill opening and closing, and swallowing. Expecting pecking behavior to develop as a pecking *pattern* can lead to the erroneous assumption that this behavior is 'innate'. The modules underlying and guiding pecking behavior do not develop in unison; full coordination is present only after birth and improves with postnatal experience. Head movement is observed as early as 3 days of embryonic age and is the earliest module of pecking behavior to be observed. The initial movement is caused by amniotic

contractions synchronized to the fetus' heartbeat. A day later, around postnatal day 4, the bill begins to open and close when the embryo's head nods. Swallowing is the last to develop and is seen only 8-9 days after bill movement. Thus pecking behavior in-utero is the result of the repeated elicitation of head-lunging and bill movement by tactile stimulation from the yolk sac. This leads to increasing ties between the two movements. The elicitation of swallowing through increased amniotic fluid pressure when the embryo opens its bill, ties swallowing to the other two modules. It is only when the chick is released from the confines of the egg and head-lunging is elicited by visual stimuli (e.g. the sight of food) that the full behavior is observed (Kuo, 1932a, c; Kuo, 1932b; Lehrman, 1953). In embryonic chicks, the head, the bill and regions involved in swallowing become increasingly coordinated through gestation; allowing the infant bird to use these body parts for the same behavior. This is only one half of the motor development story. The whole picture entails both increasing coordination and increasing differentiation of motor modules.

Study of wing movement development in chicks presents the entire picture. When embryonic chicks first start moving, the movement involves multiple body parts and is uncoordinated (this is common to many species, humans included) (Hamburger, 1963). Therefore, in young chick embryos (E9), synchronous wing-to-wing movement is as frequent as synchronous wing-to-leg movement. As fetal life progresses (E13), the wings and legs start differentiating and the two wings get more coordinated. By the very end of gestation (E17), wing–wing synchronies predominate over wing–leg synchronies (Provine, 1980; Bradley, 1999).

The primary goal of this thesis has been to re-examine primate vocal development, in light of the main tenets uncovered from the aforementioned works. These tenets being: (1) Fetal development lays the foundation for infant behaviors (2) Motor modules linked to different behaviors become increasingly differentiated; and (3) Motor modules used for the same behavior become increasingly coordinated.

In both human and nonhuman primates, sound is produced in a similar manner: Diaphragm action pushes air from the lungs and causes vibration of the vocal folds of the larynx. This sound source travels up the vocal tract, where the movements of face effectors (jaws, lips, tongue) act as a filter, passing acoustic energy at some frequencies and attenuating energy at other frequencies, according to their size and shape. The different mouth movements, along with changes in vocal fold tension and respiratory power, are what give rise to different sounding vocal signals (Ghazanfar and Rendall, 2008; Ghazanfar and Takahashi, 2014). In primates, a number of vocal signals are produced at birth. These signals are often dubbed as 'innate' (Arbib et al., 2008; Hammerschmidt and Fischer, 2008), but how they develop remains mysterious. As opposed to simply appearing de novo at birth, we hypothesized that vocalizations begin to self-organize prenatally through the interactions between spontaneous neural activity, fetal biomechanics and properties of the uterine environment (for instance, physical constraints of the uterus). To show that this is case, what is needed, is a longitudinal study using a non-invasive approach to measure discrete fetal vocalization-related movements. A longitudinal study will allow observation of the entire developmental trajectory of the behavior. A non-invasive study has limitations (for instance, it is challenging to tease apart the underlying mechanisms without manipulation) but, it will enable observation of the fetus in it's natural environment. Air is required for sound production and under normal conditions fetuses do not produce sounds (Ryder, 1943). Nonetheless, for vocalizations to be produced at birth, some actions of the vocal apparatus must be developing in-utero. Not all parts of the fetal vocal apparatus can be observed non-invasively but fetal mouth (jaw) movements can be observed. We therefore used ultrasound imaging on awake, pregnant marmoset monkeys (*Callithrix jacchus*) to track fetal movements of the mouth, from the day the mouth is first visible to the day before birth. This approach has allowed us to observe not just the start and end points of mouth movement development, but the entire trajectory, in individual pregnancies, in a non-invasive manner.

We found that mouth and head parts move together early in gestation but gradually decouple, and move independently by the end of gestation. We have evidence suggesting

that inhibition from the developing brain, differential tactile receptor distribution on the fetal body and increasing space constraint from the uterine wall all serve to organize the connections between mouth and head modules, via anti-hebbian rules of neural pruning.
Nonhuman primates are born precocial and develop at a much faster rate than humans. In addition to cry-like sounds, newborn marmosets produce calls that are structurally comparable to adult calls (Bezerra and Souto, 2008). We observed that signature features of marmoset infant calls emerge prenatally as distinct patterns of mouth movements. Late in gestation, the duration and syllable number of phee and twitter calls — calls produced by marmoset monkeys on the first postnatal day — are evident in the mouth movements of fetal marmosets. This can only be achieved by increasing coordination of the mouth musculature.

This thesis therefore shows that: (1) Aspects of vocal behaviors in marmosets have a period of prenatal development. (2) Mouth and head modules become increasingly differentiated through gestation; allowing the infant to use these regions for separate functions (mouth for feeding and vocalizing; head for orienting). (3) Mouth musculature becomes increasingly coordinated through gestation; allowing the infant to produce vocalizations on the first day of birth.

The model system we have chosen, the marmoset, is a New World primate of the family Callitrichidae. The weight of the adult marmoset averages 350 to 400 g, approximately that of a rat and about 4% of the adult weight of a rhesus macaque. Marmosets typically give birth to twin or triplet litters after a 4- to 5-month gestation, and the twins often weigh as much as 20% of the female's body weight. Littermates are hematopoietic chimeras as a result of chorionic fusion before embryonic gastrulation (Merker et al., 1988). The production of such littermates that share an amazingly integrated in-utero environment and then a similar postnatal environment offers many opportunities for studying development. Wild callitrichid primates typically live in territorial groups of 5 to 15 individuals in which only one female and one male breed. Captive housing can easily mirror this normal group structure (Tardif et al., 2011). Marmoset monkeys (like other nonhuman primates) are particularly close to humans. They have roughly the same

body plan as humans and have homologous brain structures (Burman and Rosa, 2009; de la Mothe et al., 2012). There are behavioral similarities too. They are voluble and engage in vocal turn- taking with dynamics similar to those of human conversations (Takahashi et al., 2013). Plus, they share with humans a cooperative breeding strategy (Takahashi et al., 2013). With all these commonalities, it is likely that the findings of this thesis are applicable to human development.

A similar investigation has not been attempted in human fetuses; but for all our findings we have provided analogous studies that point to homologous developmental mechanisms in humans and marmosets. Critically, the similarities we have discovered show that as unique as is it may be, our vocal behavior is a product of evolution and has it's biological roots in primate communication.

Chapter 1

Self-organization of mouth and head modules in fetal marmosets

Summary

Behaviors present at birth are possible only because of development that has happened in fetal life. In the infant, mouth and head parts are separate units that perform different functions (e.g. mouth for feeding, vocalizing; head for orienting). Using ultrasonography, we have been able to track the development of mouth and head movements in marmoset monkeys (*Callithrix jacchus*). We see that these two movements have different onset times and different developmental trajectories, and by the end of gestation both movements almost completely die out. From the pattern of decline in tested pregnancies and evidence from previous studies, we have inferred that space constraint and inhibition from the developing brain contribute to the decline in movement. By quantifying the pairing of mouth and head movements from the time the fetal face is first visible until the end of the pregnancy, we show that in marmoset monkeys, these movements initially overlap and they completely decouple only by the end of gestation. Entropy measures show that the separation allows for structure to emerge in movement behavior. From the profile of overlap decline, we have inferred that the separation of mouth and head motor modules could be driven by anti-hebbian mechanisms, triggered by the differing movement trajectories of mouth and head. In short, self-organization requires dynamic and reciprocal interactions between the brain, body and environment. Given the similarities in body plan and brain morphology of humans and other non-human primates, it is very likely that this pattern of development persists in humans.

Introduction

The fetus starts moving as soon as functioning muscles and the spinal cord appear (Hadders-Algra, 2007). These movements play an important role in the formation and

organization of the nervous system (Granmo et al., 2008; Blankenship and Feller, 2010), and are ultimately necessary for normal behavioral development (Bekoff, 2001). Early movements allow a developing organism to experience its environment within the boundaries afforded by its physical structure — its muscle connectivity, tactile receptor distribution and neural circuits (Blumberg et al., 2013). It is the dynamic interactions between the developing body and brain, and between the body and its environment, that drive the growth of muscle and neural organization (Chiel and Beer, 1997; Pfeifer et al., 2014).

Simulation studies have shown proof of this principle. For instance, Kuniyoshi and colleagues (Mori and Kuniyoshi, 2010; Mori, 2012; Yamada and Kuniyoshi, 2012) have modeled the motor development of human and zebrafish embryos. At the start line the fetal models had different sizes, musculoskeletal configurations and tactile receptor distributions, but the same neural circuits. Spontaneous movements dictated by the physical properties of the fetus led to differential neural organization, which in turn influenced movement. In both mock-ups, at first, one motor unit controlled multiple body parts (e.g. a single unit controlling the head and trunk). As development progressed these large units subdivided into more precise units (e.g. separate control of head and trunk), leading to organized, species-typical movement patterns. Pfeifer and colleagues (Marques et al., 2013) have simulated the development of reflexive motor behavior. In their model, spontaneous motor activity guided the self-organization of reflex circuitry via anti-hebbian rules of neural pruning [According to the anti-hebbian rule, when two networks are simultaneously activated but get different amounts of input, the connections between them become more inhibitory so that joint activity is discouraged and correlation is decreased (Földiák, 1990)]. They contend that two factors contribute to the development of reflexes – muscle morphology and an unperturbed environment. In both simulation studies the final behavior patterns are consistent with what has been observed in biological systems. What about the developmental trajectories? To ground these models in real life what is needed is a longitudinal study using a non-invasive approach to measure discrete fetal movements and observe how they evolve over the course of development.

Here, we have used ultrasound imaging on marmoset monkeys (*Callithrix jacchus*) to track fetal movements of the mouth and head, from the day the mouth is first visible to the day before birth. This approach has allowed us to observe not just the start and end points of motor development, but the entire trajectory in individual pregnancies, in a non-invasive manner.

Marmoset monkeys are an ideal model system for this investigation because they have roughly the same body plan as humans and have homologous brain structures (Burman and Rosa, 2009; de la Mothe et al., 2012). Simulations have highlighted the importance of body morphology for fetal motor development (Mori and Kuniyoshi, 2010; Mori, 2012; Yamada and Kuniyoshi, 2012). Given the similarity of primate bodies, it is likely that the observed developmental trajectory is characteristic of human fetal development as well. These monkeys also breed readily in captivity and regular ultrasound imaging can be done on them with no adverse effects on the fetus (Jaquish et al., 1995).

Our study is the first to analyze fetal motor development in a non-human primate species and the first longitudinal study in any non-human species. Why mouth and head regions? Let us turn to what babies can do at birth — they are very adept at noisily demanding the attention of their caregivers and are equally adept at feeding. There is more to these behaviors, but one crucial requirement for both vocalizations and suckling movements is control of the mouth musculature. Newborns can also move their heads to orient to sensory stimuli without having to simultaneously move their mouths. All of these behaviors are possible only on account of changes that have happened in fetal life. We have attempted to uncover some of these changes. Experimentally, mouth and head areas are idea candidates because they are discrete, easily observed throughout gestation and have been understudied. An added advantage of looking at these two regions is that they are probably subject to different environmental pressures. For example, change in uterine space is thought to influence motor development (Smotherman and Robinson, 1988; Bekoff, 2001). A large structure such as the head would come under increasing space constraint during gestation, versus the mouth (which likely does not or even if it does,

does so to a much lesser extent). The effect of space constraint can then be understood more fully than if we observed the development of only large or only small body parts.

The aims and predictions of our study are as follows: (1) What are the developmental trajectories of mouth and head movements? Based on studies in a number of different species (rats, ducks, pigeons, chicks, humans) we expected to see a change in movement quantity through gestation (Barron and Donald, 1941; Oppenheim, 1974; Gottlieb, 1976; Smotherman and Robinson, 1988; Roodenburg et al., 1991). In human fetuses, the first mouth movements are seen after head movement onset (DiPietro, 2008). Therefore, we also expected to find that mouth and head movements have differing developmental trajectories. (2) What is the link between mouth and head movements through gestation? In infants, the mouth and head perform different functions; therefore, the final pattern is two uncoupled body parts. The simulations of Kuniyoshi and colleagues (Mori and Kuniyoshi, 2010; Mori, 2012; Yamada and Kuniyoshi, 2012) have shown that large motor modules controlling many body parts differentiate into smaller modules with more precise control. A similar trend has been observed in chicks, where the wings and legs are coupled early in gestation and decouple before birth (Bradley, 1999). Given the anatomical proximity of mouth and head parts it is possible that they are initially coupled and decouple only later in gestation. (3) Our experiments have been observational but we have used the mechanisms uncovered by previous studies to predict the developmental profiles of mouth and head movements. Comparing these predictions with our data, we suggest that differential tactile input (Mori and Kuniyoshi, 2010; Mori, 2012; Yamada and Kuniyoshi, 2012), increasing space constraint (Smotherman and Robinson, 1988; Bekoff, 2001), and inhibition from the developing brain (Barron and Donald, 1941) must act together to drive the organization of mouth and head circuitry via anti-hebbian learning rules (Marques et al., 2013). (4) We also hypothesized that there would be less randomness in the motor behavior of late stage fetuses. This is a hallmark of self-organization.

Figure 1 summarizes the main results of our study: In a young fetus, mouth and head areas are part of the same motor module (Figure 1A, lower illustration). By late gestation,

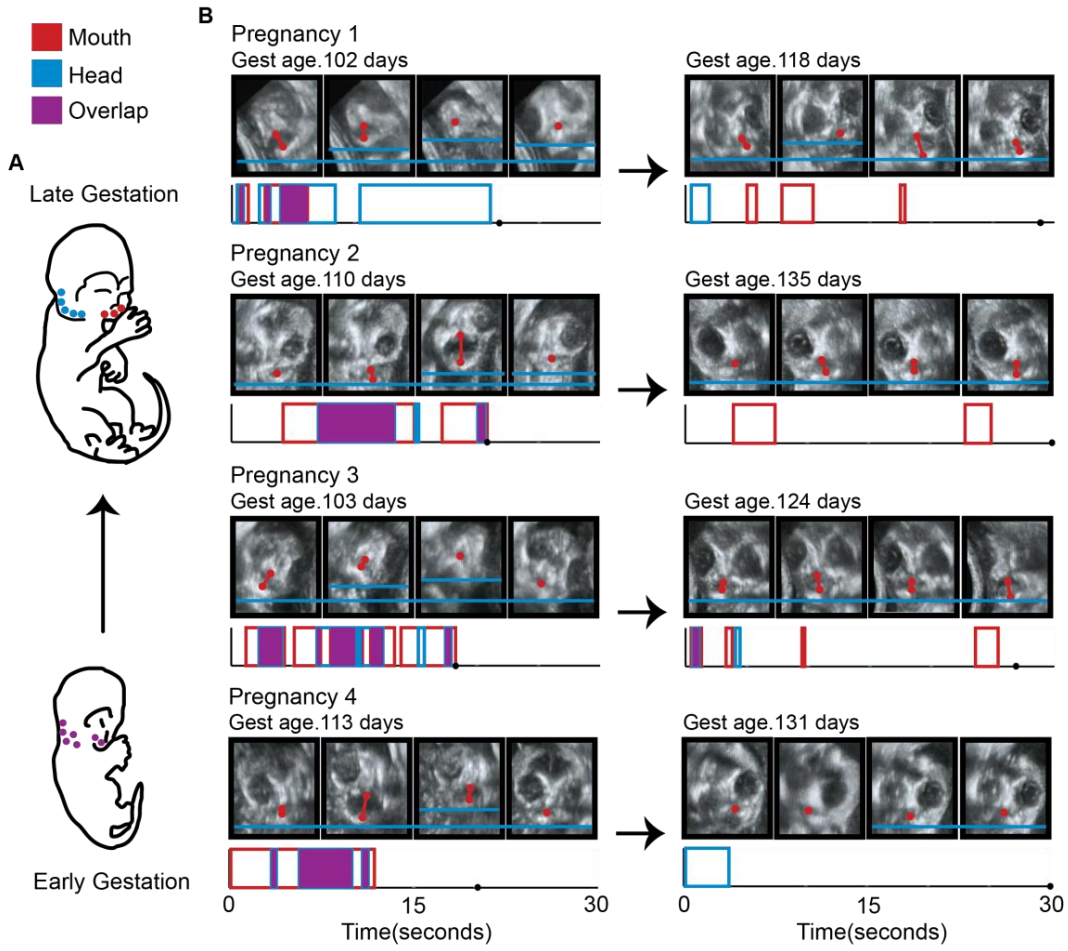


Figure 1. Exemplars of the different types of movements. Mouth movements are depicted in red, head movements in blue and overlapping mouth and head movements are shown in purple. (A) Illustration of how motor modules change through gestation. The lower panel shows an early stage fetus, where mouth and head regions are part of the same motor unit. The upper panel shows a late stage fetus. At this stage mouth and head regions have distinct motor units. (B) Exemplars that show the developmental change illustrated in (A). For each pregnancy, the early stage fetus is on the left and the late stage fetus on the right. Here, the top panel in each set shows still frames extracted from ultrasonography clips of fetal movement. The clips varied in length, with the maximum duration being 30 seconds. The red lines are used to mark mouth opening. The bottom blue lines mark the position of the head when it is at rest. The top blue lines show the position of the head after movement. The lower panel of each set is a tracking of the movement. The x-axis shows the onset and offset time of each movement. The black dot on the x-axis marks the exact length of the clip. The amplitude of movement had no bearing in our analysis so the y-axis value is the same for all movements. Mouth movements are in red and head movements are in blue. The purple shaded regions show the extent of overlap. Note that the region of overlap decreases from early gestation (panels on the left) to late gestation (panels on the right). This trend holds across all pregnancies.

this large module has differentiated, and the two areas are no longer coupled (Figure 1B, upper illustration). Our study shows behavioral proof of this change (Figure 1B). Each strip of images in Figure 1B comes from an exemplar session. Mouth openings are indicated in red and head movements in blue. Beneath the image strip is a tracing of mouth and head movements in time. The dot on the time axis indicates the length of the clip. Here too, mouth movements are in red and head movements in blue. The purple blocks mark instances of overlapping mouth and head movements. The panels on the left are from early stage fetuses and the ones on the right are from late stage fetuses. The two exemplars on each row are from the same pregnancy. The decrease in overlaps (reduced purple bocks) from early to late stage fetuses shows that mouth and head regions are coupled early in gestation and uncoupled by the end of gestation.

Results

We monitored spontaneous mouth and head movements in nine fetuses of four pregnant marmoset monkeys using ultrasonography. Of the four pregnancies, one was of a singleton, two of twins and one of quadruplets. Of the nine fetuses, seven survived (two of the quadruplets were stillborn). There were no obvious differences in the sizes of the singleton, twins, or quadruplets at birth (including the stillborn infants). Our veterinary doctor performed necropsies on the stillborn animals and concluded that both were well developed and looked to be at term. Observations of fetal mouth and head movements started on the first day the face was clearly visible (~ gestational day 95) and lasted until the day before birth (~ gestational day 146). The analysis in this paper is from observations conducted on every fourth day of gestation. In total, we have 53 sessions of testing (14 sessions from the first pregnancy and 13 sessions from each of the other three pregnancies). In each session, the ultrasound probe was positioned in a manner that enabled us to see a frontal view of the fetal face, and movements were monitored for around 15 to 45 minutes (the ultrasound was terminated if the mother became intolerant

of the procedure, hence the variability in session length). When there were multiple fetuses we picked two fetuses at random and tracked the mouth and head movements of each fetus for 15 to 20 minutes.

We then did a frame-by-frame analysis of the ultrasound videos to identify mouth and head movements. We only looked at the segments of video where mouth and head were both visible. A movement counted as a mouth movement when there was a clear separation of the upper and lower jaws (Figure 1B image panel, in red). The first video frame where the jaws separated counted as the beginning of mouth movement and the first video frame where the jaws came back together was the end point of the movement. If we lost focus of the face before the jaws came back together, the last frame of observation was marked as the end of movement. A movement counted as a head movement when there was a clear deflection of the skull from a baseline position (Figure 1B image panel, in blue — the lower line shows the baseline, the upper line marks the position of the head after movement). Similar to mouth movements, the first video frame where the head moved away from baseline position was the beginning of movement and the first video frame where the head came back to baseline was the end point of movement. If we lost focus of the fetus before the head came back to baseline position, the last frame of observation counted as the end of the movement. The amplitude of movements did not have a bearing in our analysis. Next we grouped individual movements of the same type (mouth or head) into a single movement unit if they were separated by 500ms or less. The 500ms criterion was determined based on the vocal behavior of infant marmosets, whose vocalizations consist of multiple syllables where each syllable is separated by ~500ms (DiMattina and Wang, 2006). For consistency, we applied the same criterion to the head movements. Henceforth, when we refer to mouth or head movements we are speaking of movement units. When mouth and head movements occurred simultaneously, the occurrence was scored as an overlap (Figure 1B, purple shaded regions).

As a first step, to get a sense of the developmental trajectories of mouth and head movements, we quantified the two movements through gestation. This was done to see if

the quantity of each type of movement changes with time and to compare the two trajectories. From previous studies we expected to see a change in movement quantity through gestation (Barron and Donald, 1941; Oppenheim, 1974; Gottlieb, 1976) (Smotherman and Robinson, 1988) (Roodenburg et al., 1991). Given that the onset times of the two movements differ, we also expected to see different trajectories for mouth and head movements (Narayanan et al., 1971; DiPietro, 2008).

The raster plot in Figure 2A&B show each movement as a hash; the thickness of the hash is proportional to the duration of movement. Mouth movements are in red and head movements in blue. The grey bars are used to illustrate the length of each session and the dotted black lines demarcate the tested period into four roughly equal quadrants — days 93-105 constitute the first gestational period, days 106-118 the second period, days 119-131 the third period and days 132-148 the fourth gestational period. Quadrants 1-3 each have 3 sessions from every pregnancy and the 4^{th} quadrant has 5 sessions from the singleton pregnancy and 4 sessions from each of the other pregnancies. This division was made only for purposes of explanation; it did not have a bearing on our analysis. In the raster plots, observe that the number of movements differ from quadrant to quadrant. Focusing on the mouth movements (red hashes), we see that first and last quadrant have lower movement numbers when compared to the other two quadrants. For head movements (blue hashes), the numbers of movements seem to decrease from quadrant 1 to 4.

We then quantified this change (Figure 2C, 2D). The sessions varied in length, therefore, we calculated the rate of each movement type in a session (number of movements per second). This was done by dividing the total number of movements (mouth and head separately) in each session by the total amount of time the face is seen in that session. Figure 2C shows the pooled data from all 4 pregnancies. Each dot in the figure represents the rate of movement calculated for one session. The curves were generated by first calculating the optimal degree for polynomial fitting according to Akaike's information criterion (AIC) (Akaike, 1981a). [Using AIC, for mouth movements a polynomial curve of degree 2 was found to be the best fit (second order coefficient: 0.46e-03); for

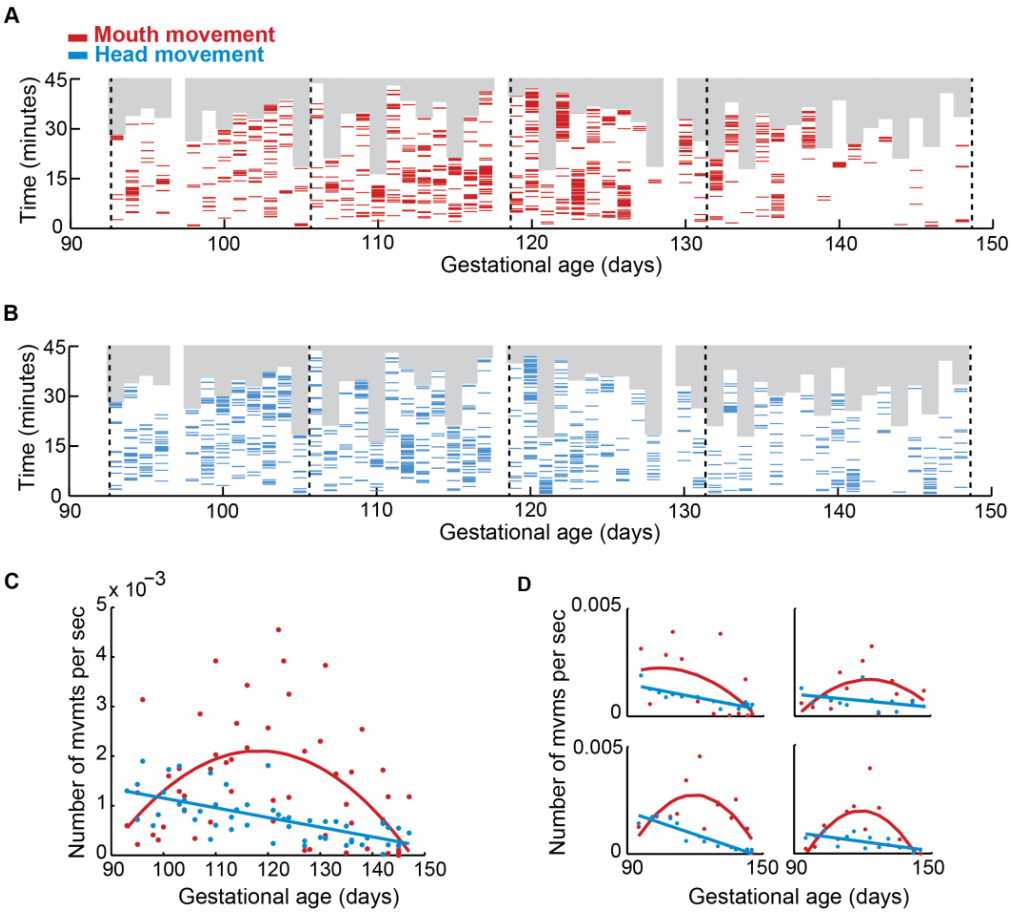


Figure 2. Trajectory of mouth and head movement rates through gestation. (A) Raster plot with mouth movements. The x-axis represents estimated gestational age in days and the y-axis denotes time. Each column represents a single session with the dotted black lines dividing the testing period into four roughly equal quadrants. Each red hash represents a single movement unit and the thickness of the hash is proportional to the duration of the movement. Observe that the number of movements is higher in the second and third quadrant compared to the first and last. (A) Raster plot with all head movements. The x-axis represents estimated gestational age in days and the y-axis denotes time. Each column represents a single session with the dotted black lines dividing the testing period into four roughly equal quadrants. Each blue hash represents a single movement unit and the thickness of the hash is proportional to the duration of the movement. Observe that the numbers of movements decrease from the first to the fourth quadrant. (B) Quantification of data represented in the raster plot. The x-axis shows gestational age in days. The y-axis represents the number of movements observed per second. Each dot represents the rate of movement in one session (red for mouth, blue for head). The curves were generated by first finding the optimal degree for polynomial fitting according to Akaike's information criterion (AIC), on the pooled data [mouth movement: degree; head movement: degree 1]. This optimal degree was then used to generate the rate profiles of head and mouth movements for all the graphs (pooled and individual). (C) Trajectory of mouth and head movement rates through gestation, for each pregnancy. The legend is the same as in (B)

head movements a polynomial curve of degree 1 was found to be the best fit (first order coefficient: 0.0001e-09]. As seen in the figure, mouth and head movements follow different developmental trajectories. Mouth movements show an inverse U-shaped profile: an increase in rate of movement in the second and third gestational period, followed by a decline in the fourth gestational period. This profile is similar to what is seen with many types of fetal movements, including, jaw movements of human fetuses (Roodenburg et al., 1991). On the other hand, in our time of observation, head movements show a linear trend, with the highest rate of movement in the first gestational period and a decline in movement rate through gestation. In the developing fetus, head movements appear before mouth movements; we are probably catching the tail end of the head trajectory (Narayanan et al., 1971; DiPietro, 2008). To see how well these trends held across individual pregnancies, the dataset was partitioned by pregnancy and polynomial curves of the same degree as the pooled data were fitted for each. The developmental trends seen with the pooled results held across all four pregnancies (Figure 2D). [First pregnancy: mouth movements — second order coefficient = 0.43; head movements — first order coefficient: = 0.016e-03; sample size = 14 sessions. Second pregnancy: mouth movements — second order coefficient = 0.04; head movements — first order coefficient: = 0.001; sample size = 13 sessions. Third pregnancy: mouth movements — second order coefficient = 0.04; head movements — first order coefficient: = 0.07e-03; sample size = 13 sessions. Fourth pregnancy: mouth movements — second order coefficient = 0.14; head movements — first order coefficient: = 0.006; sample size = 13 sessions.]

We can conclude that the quantities of mouth and head movements change over fetal life and that they follow different developmental trajectories. The observed profiles are in keeping with results from previous work, on various types of movements, in a number of animal species (rats, ducks, pigeons, chicks, humans) (Barron and Donald, 1941; Oppenheim, 1974; Gottlieb, 1976; Smotherman and Robinson, 1988; Roodenburg et al., 1991).

As seen in Figure 2, by the end of gestation, there is very little movement of any type. The reason for decline in movement is still unknown but there has been some speculation on the matter. One theory states that movement decline is due to decrease in uterine space, as the fetus gets larger (Smotherman and Robinson, 1988; Bekoff, 2001). If this is the case, we can expect head movements to start declining earlier than mouth movements; which is what we see. We agree that decrease in uterine space must contribute to decrease in movement but contend that the decline cannot be explained solely by space constraint. Our reasons are as follows.

In this study data was collected from four different pregnancies — one singleton, two sets of twins and one set of quadruplets. If space constraint was the sole reason for decline, we should see that the starts of decline differ based on the number of fetuses present in the uterine environment. Figure 3 shows our results partitioned by pregnancy – the singleton pregnancy is in pink; the first twin pregnancy in blue; the quadruplet pregnancy in green; the second twin pregnancy in orange. Figure 3A shows the biparietal distance of fetal skull. For the pregnancies with multiple fetuses, each point represents the average biparietal distance of the two infants imaged in that session. The fitted curves were generated using the same method used in previous figures [Using AIC, First pregnancy: first order coefficient = -0.14. Second pregnancy: first order coefficient = 0.16. Third pregnancy: fourth order coefficient: = -5.46e-06. Fourth pregnancy: fourth order coefficient: = -1.55e-06]. The skull size trajectories overlap, showing that there is very little variation in the skull size of individual fetuses. The trajectories match up with previous experiments (Jaquish et al., 1995). Figure 3B shows biparietal distances of the same fetuses on day 1 of postnatal life, including those of the two stillborn fetuses from the quadruplet pregnancy. Here too, there is not much difference between individuals. Therefore, if there are more individuals, there should be less space in-utero. Figure 3C shows the mouth movement profiles for the individual pregnancies. If the space constraint theory was true, the decline should start earliest in the quadruplet pregnancy and last in the singleton, and in Figure 3D we should see flat lines connecting columns 1 and 2. This is not the case. In fact, mouth movement decline begins earliest in the singleton, around 20 days earlier, and is about the same for the rest of the pregnancies. It

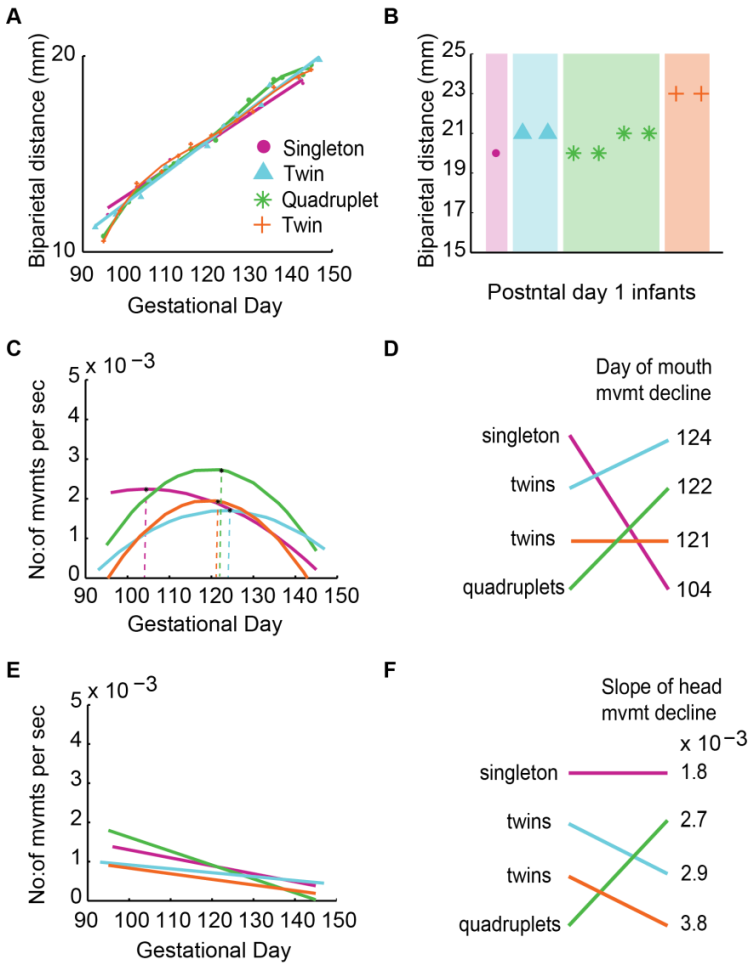


Figure 3. Space constraint does not explain the differences seen in the decline of movement rates and overlap trajectories for each pregnancy. The singleton pregnancy is in pink, the first twin pregnancy in blue, the quadruplet pregnancy in green and the second twin pregnancy in orange. (A) Biparietal distance (BPD) of fetal skull. The x-axis shows gestational age in days. The y-axis represents the fetal skull BPD in millimeters. For the pregnancies with multiple fetuses, each point represents the average BPD of the two infants that were imaged per session. (B) Biparietal distance (BPD) of infant skull at P1. The x-axis indicates the pregnancy and the y-axis shows infant skull BPD in millimeters. (C) Graph showing the profiles of mouth movement rates for each pregnancy. The x-axis shows gestational age and the y-axis shows the number of mouth movements per second. The black asterisk marks the peak of each profile and the dotted line indicates the corresponding gestational age. (D) Graph relating the gestational day of peak mouth movement (after this day the mouth movements decline) to the number of fetuses in a pregnancy. (E) Graph showing the profiles of head movement rates for each pregnancy. The x-axis shows gestational age and the y-axis shows the number of head movements per second. (F) Graph relating the head movement decline to the number of fetuses in a pregnancy.

could be that mouth movements are not affected by decreasing space. Figure 3E shows the head movement profile of each pregnancy. We were interested in the relationship between mouth and head movements so our observation of the head only began at mouth movement onset. We can however, look at the slope of head movement decline. If space constraint was solely responsible for decline, we should see the steepest decline for the quadruplet pregnancy and the gentlest decline for the singleton; in Figure 3F we should see flat lines connecting columns 1 and 2. This is not the case. Given the trends of our data, it is highly likely that movement decline is not solely due to decrease in uterine space with increasing fetal size.

Next we looked at the relationship between mouth and head areas. Infants have independent control of these regions. In fetuses, based on the simulations of Kuniyoshi and colleagues, the expectation was that a large motor module controlling both mouth and head parts would differentiate into smaller modules with separate control of the body parts (Mori and Kuniyoshi, 2010; Mori, 2012; Yamada and Kuniyoshi, 2012). Still, this is not a given, mouth and head movements are anatomically proximate so could be part of the same initial module but they have differing developmental trajectories — will they then start off as differentiated units or do they start off coupled and differentiate into independent movements?

To test between these two possibilities we looked at the instances of overlapping mouth and head movements. At first pass we did not consider the order of movements (it did not matter if mouth movements came first or if the head came first). The raster plot in Figure 4A shows each instance of overlap as a purple hash; the thickness of the hash is proportional to the duration of overlap. As in the raster plots of Figure 2, grey bars are used to illustrate the length of each session and the dotted lines demarcate the four gestational periods. Comparing the number of hashes in the four quadrants, each quadrant seems to have fewer instances of overlap than the previous one. This is particularly apparent in the third and fourth quadrants. We quantified this decrease by calculating the total number of mouth plus head movements in a session and then determining what percentage of this number is composed of overlap instances. Figure 4B shows the pooled

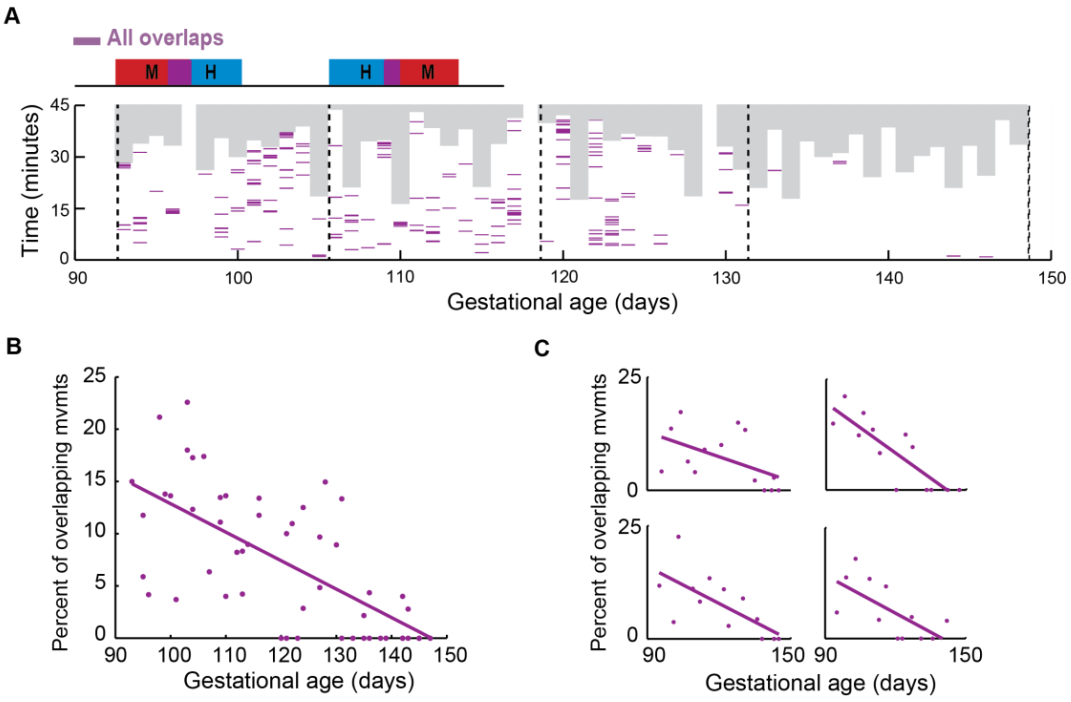


Figure 4. Overlap of mouth and head movements declines through gestation. (A) Raster plot with all the segments of mouth and head movement overlap (the order of the movements did not matter). The x-axis represents gestational age in days and the y-axis denotes time. Each column represents a single session with the dotted black lines dividing the testing period into four roughly equal quadrants. Observe that the amount of overlap decreases through gestation. (B) Quantification of data represented in the raster plot – percentage of overlapping movements per session across all pregnancies. The x-axis shows estimated fetal gestational age in days. The y-axis represents the percentage of overlapping movements. Each dot represents the percentage of overlap in one session. The curve was generated by first finding the optimal degree for polynomial fitting according to Akaike's information criterion (AIC), on the pooled data [a polynomial curve of degree 1 was found to be the best fit]. This optimal degree was then used to generate the overlap profile for all graphs (pooled and individual). (C) Trajectory of mouth and head movement overlaps through gestation, for each pregnancy. The legend is the same as in (B)

data from all 4 pregnancies. Each dot in the figure represents the percentage of overlap in one session. The curve was generated by first calculating the optimal degree for polynomial fitting according to Akaike's information criterion (AIC) (Akaike, 1981a) [Using AIC, a polynomial curve of degree 1 was found to be the best fit (first order coefficient: -0.003)]. As seen in the figure, overlap steadily decreases through gestation and by the last gestational period, overlaps have almost completely died out. To see how well this pattern held across individual pregnancies, the dataset was partitioned by pregnancy and a polynomial curve of the same degree as the pooled data was fitted for each. The developmental trend seen with the pooled results held across all four pregnancies (Figure 4C) [First pregnancy: first order coefficient: = -0.002; sample size = 14 sessions. Second pregnancy: first order coefficient: = -0.004; sample size = 13 sessions. Third pregnancy: first order coefficient: = -0.003; sample size = 13 sessions. Fourth pregnancy: first order coefficient: -0.003; sample size = 13 sessions.]. This is indication that mouth and head movements are linked early gestation and separate over time, but from the analyses shown so far, making a definitive statement on mouth and head differentiation is premature. Our next two analyses confirm that what we see is indeed a decoupling of the two movement types.

Figure 4 shows mouth and head overlap declining through fetal life. However, the total number of movements also decline (Figure 2). The decline in overlap could simply be due to the decline in total movement. To check if this was the case, we did a permutation test and contrasted the resulting trajectory with the actual trajectory (Figure 5A). For the permutation test, we took each session and independently shuffled the head movements and mouth movements, keeping the duration and latency distributions intact. We then calculated percentage of overlap, same as with the dataset. The result of such a test shows what the trend would look like if mouth and head movements were not linked. In Figure 5A, each grey dot shows the median percentage overlap for one session, after 1000 permutation. The permutation trend (grey line) was generated using the same procedure as in other figures [Using AIC, a polynomial curve of degree 1 was found to be the best fit (first order coefficient: -6.72e-05)]. The grey shaded region is the 95% confidence interval for the permuted data (the yellow dots and green dots show the upper 97.5 and

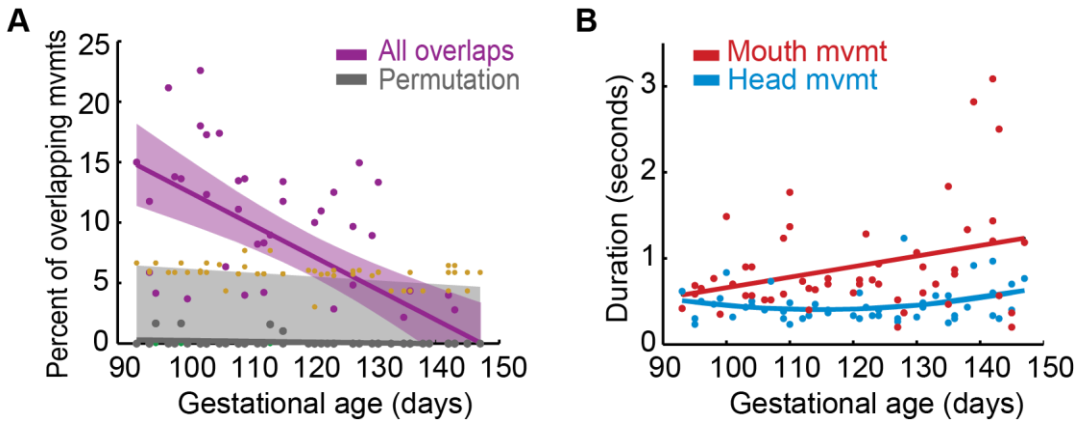


Figure 5. Overlap decrease not explained by movement decline or change in movement duration. (A) Graph contrasting the actual decline in overlap (all movements, irrespective of order) with the overlap calculated after permutation. [Permutation test: For each session, mouth and head movements were independently shuffled, keeping the duration and latency distributions intact. The reshuffled movements were then combined and amount of overlap was calculated.] The x-axis shows gestational age in days. The y-axis represents the percentage of overlapping movements. The plot for the experimental data is same as that in Figure 3. The shaded region for the data (purple) denotes the confidence interval for the fit. Each large grey dot represents the percentage of overlap in one session. The curve was generated by first finding the optimal degree for polynomial fitting according to Akaike's information criterion (AIC) [a polynomial curve of degree 1 was found to be the best fit]. The shaded region for the permutation test (grey) denotes the 95% confidence interval for overlaps calculated on the permuted data set. The yellow dots and green dots show the upper 97.5 and lower 2.5 percentile points, respectively, for each session. (B) Change in duration through gestation. The x-axis shows gestational age in days. The y-axis represents the duration of mouth and head movements in seconds. Each dot in the figure depicts the median duration of mouth movement (red) or head movement (blue) for one session. The curves were generated in the same manner as in the previous figures [a polynomial curve of degree 1 was found to be the best fit for mouth duration and a polynomial curve of degree 2 was found to be the best fit for head duration].

lower 2.5 percentile points, respectively, for each session). The purple line shows the actual results (same as Figure 4C). The purple shaded region represents the 95% confidence interval for the fit. The permutation test generated an almost flat profile compared to the overlap decline seen in our dataset. The difference between the permuted and actual dataset is significant at the beginning of gestation but not so in the third and fourth quadrants. This proves that the decline in overlap is not simply due to the decline in movement quantity.

Another (albeit slim) possibility could be that movement duration decreases through gestation and therefore, overlap decreases. To test if this was the case we calculated the median duration of mouth movements and head movements for each session and looked at the change in duration over time (Figure 5B). Each dot in the figure represents the median duration of movement in one session (red for mouth and blue for head). The trend was fitted using the same procedure used in all the previous figures. [Using AIC, a polynomial curve of degree 1 was found to be the best fit for mouth durations (first order coefficient: 0.012); a polynomial curve of degree 2 was found to be the best fit for head durations (second order coefficient: 2.15e-04)] The durations of mouth movements increase with time and the durations of head movements remain fairly steady through gestation. The decrease in overlap is not explained by a decrease in the duration of mouth or head movements.

We can now conclusively state: Early in gestation mouth and head movements are coupled, as gestation progresses they become increasingly independent and by the end of gestation they are two distinct movement patterns. This finding is in keeping with the results from simulation tests and experimental data from developing chicks (Bradley, 1999; Mori and Kuniyoshi, 2010; Mori, 2012; Yamada and Kuniyoshi, 2012).

We contend that this change is brought about by anti-hebbian pruning of the neural network (Földiák, 1990), driven by the difference in mouth and head movement quantities. Figure 6A illustrates our hypothesis. The red circles represent the mouth and the blue circles represent the head. The small arrows (over the circles) represent

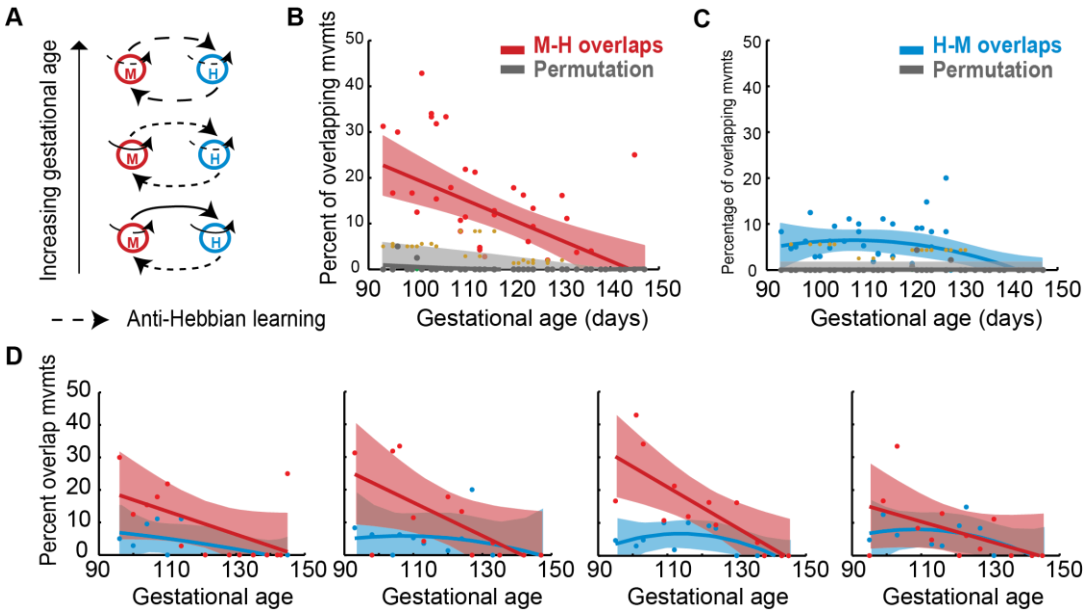


Figure 6. Anti-Hebbian rules of neural pruning can be used to explain decoupling of mouth and head movements. (A) Schematic showing increasing space constraint and its effect on the neural connections between mouth and head areas. The red circles represent the mouth and the blue circles represent the head. The small arrows represent movement. When they are dashed, movement is less. The large arrows represent the connections between the mouth and head regions. Connections weakened through anti-hebbian learning are illustrated using the large dashed arrows. The lowest panel (early in gestation) depicts a time when neither the mouth nor the head is under constraint. The middle panel depicts a time when head movement is restricted due to space constraint but not mouth movement. The top panel (late in gestation) depicts a time when both mouth and head movement is restricted. (B) Percentage of mouth movements followed by an overlapping head movement. The x-axis has gestational days and the y-axis represents the percentage of coupled mouth-head movements. The result from the experimental data is in red and the result from the permutation test is in grey. Each dot represents the results from one session (yellow dots – upper 97.5 points of permuted date; green dots – lower 2.5 percentile points of permuted data). The polynomial fit and confidence interval were calculated using the method described for Figure 4C (C) Percentage of head movements followed by an overlapping mouth movement. The x-axis has gestational days and the y-axis represents the percentage of coupled head-mouth movements. The result from the experimental data is in blue and the result from the permutation test is in grey. Each dot represents the results from one session (yellow dots – upper 97.5 points of permuted data; green dots – lower 2.5 percentile points of permuted data). The polynomial fit and confidence interval was calculated using the method described for Figure 4C (D) Trajectory of mouth and head movement overlaps through gestation, for each pregnancy. The legend is the same as in (B) & (C).

movement. When they are dashed, movement is less. The large arrows represent the connections between the mouth and head regions. Weak connections are illustrated using the large dashed arrows. Figures 6B&C show the experimental data that support our hypothesis. In Figure 6B, each red dot represents the percentage of mouth movements that are accompanied by an overlapping head movement for a single session (M-H overlap). The trajectory was fitted using the same procedure as in all other figures [Using AIC, a polynomial curve of degree 1 was found to be the best fit (first order coefficient: -0.004)]. The red shaded region is the 95% confidence interval of the fit. Each grey dot represents the percentage of mouth movements that are accompanied by an overlapping head movement after the session has been permutated. The permutation test was done using the same procedure as in Figure 5A. The curve was fitted using the same procedure as in previous figures [Using AIC, a polynomial curve of degree 2 was found to be the best fit (first order coefficient: 8.24e-06)]. The grey shaded region represents the 95% confidence interval of the permuted data (the yellow dots and green dots show the upper 97.5 and lower 2.5 percentile points, respectively, for each session). In Figure 6C, each blue dot represents the percentage of head movements that are accompanied by an overlapping mouth movement for a single session (H-M overlap). The trajectory was fitted using the same procedure as in all other figures [Using AIC, a polynomial curve of degree 2 was found to be the best fit (second order coefficient: -5.58e-05)]. The blue shaded region is the 95% confidence interval of the fit. Each grey dot represents the percentage of head movements that are accompanied by an overlapping mouth movement after the session has been permutated (the yellow dots and green dots show the upper 97.5 and lower 2.5 percentile points, respectively, for each session). The curve was fitted using the same procedure as above [Using AIC, a polynomial curve of degree 0 was found to be the best fit (zero order coefficient: 0.001)]. The grey shaded region represents the 95% confidence interval of the permuted data.

Our hypothesis: The starting point is a network with linked mouth and head areas. Early in gestation there is less mouth movement than head movement (Figure 2B). Therefore, every time the mouth moves, the head can move but when the head moves the mouth does not move as much. The connection between the mouth and head will thus be

stronger than the connection between the head and mouth (Figure 6A, lowest panel). This explains the larger percentages of M-H overlap, compared to the percentages of H-M overlap (Figures 6B&C, early gestational age). Mid gestation, the mouth movements increase in number and the head movements start decreasing (Figure 2B). Now, when the mouth moves, the head does not move as much but when the head moves the mouth can move. The connection between the mouth and head will weaken but the connection between the head and mouth will not further weaken (Figure 6A, middle panel). This explains the percentages of M-H overlap decreasing, and the percentages of H-M overlap remaining about the same (Figures 6B&C, mid gestational age). By late gestation, both movements are on the decline and the connections are already weak (Figure 2B). The few times a movement of either type occurs, it is rarely accompanied by an overlap. The connections on both sides will further weaken (Figure6, top panel), causing a decline in all overlaps (Figures 6B&C, late gestational age). Figure 6D shows that the trend we see hold across individual pregnancies. The methods and legend are the same as those for Figures 6B&C. [First pregnancy: M-H overlaps — first order coefficient = -0.003; H-M overlaps —second order coefficient: = -9.91e-06. Second pregnancy: M-H overlaps — first order coefficient = -0.005; H-M overlaps — second order coefficient: = 2.78e-06. Third pregnancy: M-H overlaps — first order coefficient = -0.006; H-M overlaps — second order coefficient: = -3.02e-05. Fourth pregnancy: M-H overlaps — first order coefficient = -0.003; H-M overlaps — second order coefficient: = -3.76e-05].

One of the hallmarks of development is the appearance of structure in movement. Young fetuses have chaotic action patterns – they have uncoordinated movements that involve many body parts and isolated jerky movements that are not part of any action pattern. With age, fetuses establish relatively organized actions patterns, reminiscent of movements at birth (Bekoff and Lau, 1980; Smotherman and Robinson, 1988). We have observed the relationship between mouth and head movements right from the onset of mouth movement, to the very end of gestation. We can therefore directly test if, for these two behaviors, there is an increase in structure through fetal life.

Mouth and head movements occur in five different ways: independent mouth movement

(State1), independent head movement (State2), mouth movement followed by an overlapping head movement (State3), head movement followed by an overlapping mouth movement (State4), and mouth-head movements with synchronous onset (State5). Figure 7A shows the proportion of states in every session, for all pregnancies. As in Figures 2 and 4, the dotted lines demarcate the four gestational quadrants. In this figure we see two distinct but related patterns (1) The variability within a single session decreases. In the first three quadrants, far more states are present in one session. By the last quadrant, for the majority of sessions, States 3-5 have disappeared and only States1&2 (independent mouth movement; independent head movement) exist. (2) The number of behavioral states decreases through fetal life. The distribution seen in the first few sessions is different from that seen in the sessions just before birth.

These changes were quantified in two ways. First, Shannon entropy was used to measure the unpredictability/randomness within each session (Figure 7B). The top value on the y-axis – 2.32 BITS – is the maximum entropy for a behavior with 5 possible states. Each dot in the figure represents the entropy measure for one session. The curve was fitted using the same procedure as in all other figures. [Using AIC, a polynomial curve of degree 3 was found to be the best fit (third order coefficient: 1.85e-05)]. We see that entropy decreases through gestation and steadies at around 0.8 BITS. The decrease in entropy measure shows that with increasing age, there is increasing predictability/structure within a single session. Next, we performed a Kullback-Leibler divergence test to quantify change through fetal life (Figure 7C). For this test, the state distribution on the earliest testing day was compared with the state distributions of every testing day. The dots on the graph represent the divergence estimates. The trend line was fitted using the same procedure as in all other figures. [Using AIC, a polynomial curve of degree 1 was found to be the best fit (first order coefficient: 0.0392)]. The linear fit with a positive slope indicates that with increasing gestational age, behavior (with respect to mouth and head movement) becomes increasingly different from the behavior at mouth movement onset. Lastly, we partitioned the results by pregnancy. Figure 7D shows the state diagrams for exemplar sessions from the four quadrants, for every pregnancy. The widths of the arrows reflect the transition frequencies between states. The first row

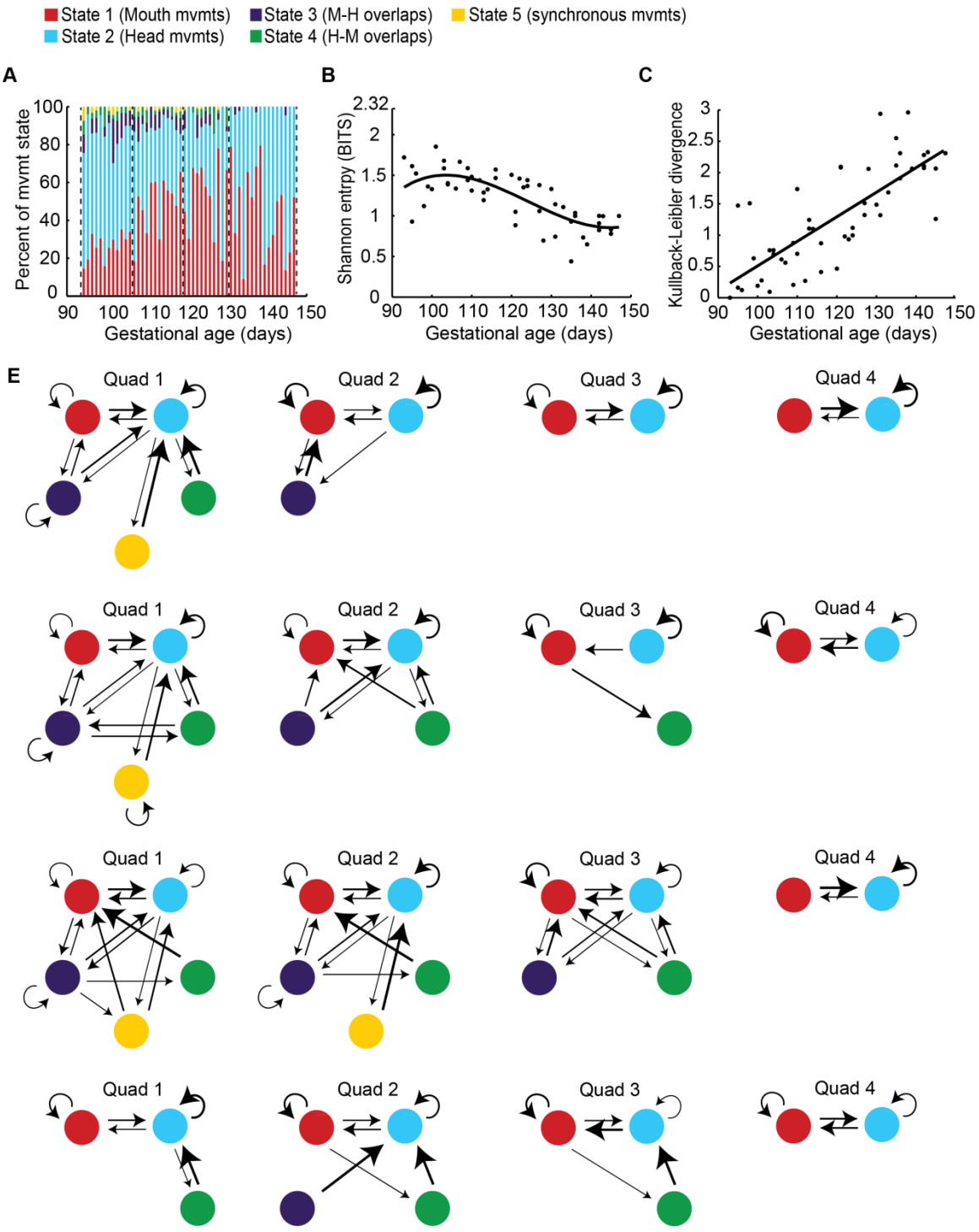


Figure 7. Increase in 'order' through gestation. Independent mouth movements have been labeled as State1 and are illustrated in red; independent head movements have been labeled as State2 and are illustrated in blue; coupled mouth followed by head movements have been labeled as State3 and are illustrated in purple; coupled head followed by mouth

movements have been labeled as State4 and are illustrated in green; mouth and head movements with synchronous onsets have been labeled as State5 and are illustrated in yellow. (A) The proportion of each behavioral state per session. The x-axis shows gestational age in days. The y-axis represents the proportion of each state. Note that the number of states decreases through gestation and in the last quadrant only State1 and State2 are consistently seen. (B) Entropy measure through gestation. The x-axis shows gestational age in days. The y-axis represents Shannon's entropy in BITS. Each dot represents the entropy measure for a single session. The curve was fitted using the same method as above [Using AIC, a polynomial curve of degree 3 was found to be the best fit] (C) Kullback-Leibler divergence to quantify the difference between the distribution of states on the first tested gestational day and the distribution on all other tested gestational days. The x-axis shows gestational day and the y-axis shows the KL divergence. Each dot represents the divergence measure for one session. The curve was generated by first finding the optimal degree for polynomial fitting according to Akaike's information criterion (AIC); this optimal degree was then used to generate the linear fit [a polynomial curve of degree 1 was found to be the best fit]. (D) Markov chain for one exemplar session from each of the four quadrants. The widths of the arrows are proportional to the transition frequencies between states. The first row consists of exemplars from the singleton pregnancy; the second row from the first twin pregnancy; the third row from the quadruplet pregnancy and the last row consists of exemplars from the second twin pregnancy.

consists of exemplars from the singleton pregnancy; the second row from the first twin pregnancy; the third row from the quadruplet pregnancy and the last row consists of exemplars from the second twin pregnancy. Notice that the first 3 quadrants differ among the pregnancies but by the fourth quadrant a consistent pattern of independent mouth and head movements has been reached.

To conclude: In the developing fetus, mouth and head movements have different onset times (Figure 2). Early in gestation mouth and head areas are controlled by the same motor module; as gestation progresses these areas uncouple — as evidenced by a decrease in overlap (Figure 4). This self-organization is achieved via anti-hebbian pruning of the controlling motor module (Figure 7); a mechanism put into action by the differing movement trajectories of mouth and head areas (Figure 2).

Discussion

The abilities an infant possesses on the day of birth have to (at least in part) have developed in-utero. As fetal life progresses order emerges from early haphazard movements (Barron and Donald, 1941; Hamburger, 1963; Prechtl, 2001). Some muscle groups become increasingly coupled with time (e.g. ipsilateral limb muscles). Other muscle groups, initially part of the same motor unit, become increasingly differentiated with time (e.g. contralateral limb muscles; wings and legs of chicks). These regions can then be used to perform separate functions in infant life. Proof of this can be seen in both simulation and experimental studies (Barron and Donald, 1941; Hamburger, 1963; Bradley, 1999; Yamada and Kuniyoshi, 2012). We looked at two anatomically close regions that perform separate functions in infants — the mouth and head regions. We see that these two movements have different developmental timelines but are linked early in gestation. Late in gestation the two movements uncouple (Figure 1).

Following their first appearance mouth movements gradually increase in quantity and after hitting a peak in the second quadrant of gestation the movements decrease in

quantity, almost completely dying out by the end of gestation (Figure 2, red). In the case of head movements we see that the quantity is highest at the beginning and declines in a linear fashion through the period of testing, also reaching very low levels by the end of gestation (though not quite as low as the mouth movements) (Figure 2, blue). Head movements start earlier than mouth movements (Narayanan et al., 1971; DiPietro, 2008). The quantity of head movements likely follows the same inverse U-shaped trajectory as the mouth movements but we are only observing a segment of their development. This inverse U-shaped trajectory is fairly universal: the activity levels of many different species (rats, ducks, pigeons, chicks, humans) follow this trajectory (Barron and Donald, 1941; Oppenheim, 1974 {Smotherman, 1988 #44; Gottlieb, 1976) (Roodenburg et al., 1991). Most of these studies have looked at total activity levels but the quantities of specific movement types also show this trend. For instance, human eye saccades, respiration movements and jaw movements show this trend (Roodenburg et al., 1991).

The relationship between these two movements also changes through fetal life. In marmoset fetuses, early in development mouth and head movements show many instances of overlap; this link almost completely disappears by late gestation (Figure 4). Comparing overlap decline to the profile generated by permutation of the dataset (Figure 5A) we have determined that the overlap decline is not related to the general decline in movement. Additionally, the decline in overlap is not due to a decrease in the duration of mouth and head movements (Figure 5B). We have concluded that mouth and head parts are controlled by the same motor module in young fetuses and, that this large module gradually differentiates into smaller modules allowing for separate control of mouth and head parts. What we see here is similar to what Kuniyoshi and colleagues show in their simulations (Mori and Kuniyoshi, 2010; Mori, 2012; Yamada and Kuniyoshi, 2012). Ties between the mouth and head region have also been observed in aborted human fetuses (Hooker, 1936; Fitzgerald and Windle, 1942). Hooker and colleagues found that a light touch to the perioral region elicits contractions of the neck muscles. Fitzgerald and Windle have found that stimulation of the maxillary region of the mouth causes contralateral head flexions (Fitzgerald and Windle, 1942).

What could be the mechanisms responsible for the observed patterns? First, the general profiles of mouth and head movement rates might be the same but the quantities of movements differ. Mouth and head movements could have been initiated in multiple ways and we did not differentiate between these different types of movements. The movements could have been (1) myogenic in nature — produced by the contraction of muscle fibers in the absence of any nervous input, (2) neurogeneic in nature — produced through neuro-muscular mechanisms that are self-generated or induced by endogenous stimuli (such as agents in the circulation) or, (3) reflexive in nature — activated by the developing sensory system (movement classification taken from: {Hamburger, 1963 #26}). Mouth and head movements could have differing quantities of each type of movement. For instance, the mouth regions of animals have a larger number of tactile receptors in comparison to other parts of the body (Grossman et al., 1965). Different levels of proprioceptive feedback and different levels of tactile stimulation [human fetuses are known to touch their face with their hands (Prechtl, 2001)] received by mouth and head parts could undoubtedly serve to produce differing numbers of reflexive movements and therefore total movement. In fact, Kuniyoshi and colleagues (Mori and Kuniyoshi, 2010; Mori, 2012; Yamada and Kuniyoshi, 2012) have found that differential tactile feedback is a driving force behind the differentiation of large muscle units into smaller ones. Therefore, their differing quantities could drive the uncoupling of mouth and head movements.

Mouth and head movements decline by the end of gestation. The reason for decline is still unknown but there has been some speculation on the matter. Two theories have been put forward: One theory proposes that the decline in movement is due to increasing space constraint, as the fetus gets larger (Hayat et al., 2011; Nowlan, 2015). The other postulates that space constraint cannot explain the decline in movement, development of inhibitory neural mechanisms causes the decline (Barron and Donald, 1941; Prechtl et al., 1979). Our observations of fetal marmoset development can be explained only if both of these theories hold true.

First, we see that head movement starts declining earlier than mouth movement. The head being larger than the mouth, this is what we would expect if space constraint caused decline in movement. Studies on limb movements in chicks show that the degree of limb movement is restricted once the fetus gets larger and space is tight (Sharp et al., 1999). It is very likely that in a similar manner, a large structure such as the head is under space constraint with increasing fetal age. However, when looking at the movement trajectories of individual pregnancies the space constraint theory does not entirely hold (Figure 3). Of the four pregnancies we observed, one was of a singleton, two of twins and one a quadruplet. Fetal skull sizes in all pregnancies were very similar (Figure 3A) and the animals on the day of birth had comparable skull sizes (Figure 73B). If space constraint alone was responsible for movement decline we should see movement decline start earliest in the quadruplet pregnancy and last in the singleton pregnancy. This is not the observed pattern. In fact, movement decline starts earliest in the singleton, around 20 days earlier (Figure 3 C,D,E&F). Also, at the end of gestation, despite the decrease in total movement, there are more head movements than mouth movements (Figure 6A). Further, Smotherman and colleagues (Smotherman and Robinson, 1988) have compared the activity of fetal rats in three different environments: the uterine environment, a condition of slight restraint (fetus externalized from the uterus with the extra-embryonic membranes intact) and a condition of no restraint (in a bath with all membranes removed). The activity levels of the embryo increased with decrease in restraint but the trajectory of activity had the familiar inverse U-shaped profile in all three cases. Lastly, a decline in movement quantity is seen for slight movements: mouth movements, eye saccades and even respiratory movements (Roodenburg et al., 1991). It is hard to argue that these movements, particularly the eye saccades and respiratory movements, decrease solely due to space constraints. Adding all the pieces of evidence together, it is very likely that decreasing space is not the only factor contributing to the decline movement at the end of gestation.

The other theory is that the decline in movement quantity is due to inhibition from the brain. This theory was proposed by Barron and colleagues based on work in fetal sheep (Barron and Donald, 1941). The details of the theory: As brain development advances,

there is an increasing level of inhibition, which eventually causes the fetus to enter a second phase of inactivity. The brain develops in an inferior-to-superior direction. First the spinal cord and lower medulla develop, then structures in the midbrain become active and then the basal ganglia is formed. The cortex seems active in some species but not in others. For instance, in rabbits the motor cortex is not excitable before birth but in sheep and monkeys it is excitable (Barron and Donald, 1941). Barron proposes that movement inhibition at the end gestation is probably due to the basal ganglia. There have been other experiments examining supra-spinal influences on motility. Viktor Hambruger (Decker and Hamburger, 1967) did a series of experiments where he performed surgical extirpations of different regions of the embryonic chick brain. The chicks with ablated midbrain regions showed a linear increase in movement through the length of gestation. This finding supports Barron's theory of inhibition from the basal ganglia. Clinical work on premature infants shows that despite being removed from the confines of the uterus these infants have decreased movement quantity similar to late stage fetuses (Prechtl et al., 1979). Therefore, inhibition from the developing brain (and not space constraint) is thought to be the cause of movement decline in human infants as well. In developing fetuses transitory brain structures called the ganglionic eminence (GE) provide the bulk of cortical inhibitory GABAergic neurons and contributes to the formation of the basal ganglia (Wichterle et al., 2001; Rallu et al., 2002; Encha-Razavi and Sonigo, 2003). In the marmoset fetus the GE is visible in the 12^{th} week of gestation and disappears in the 16^{th} week of gestation (Hikishima et al., 2013). In our dataset the decline in mouth movement starts at the end of the 16^{th} week of gestation (118 gestational days), on average. The lining up of the two timings suggests that the decline in movement could be related to the maturation of inhibitory processes in the brain.

We suggest that tactile receptor distribution, space constraint and, inhibition from the developing brain all contribute toward the mouth and head movement quantity profiles. The values of each of these contributors are different for the mouth and head regions; therefore, their developmental trajectories are different.

Increasing space constraint and inhibitory neural mechanisms can also be used to explain the uncoupling of mouth and head regions. The starting point is a coupled network for mouth and head but since the two parts have differing movement profiles (Figure 2), the network gets differing amounts of input from the two regions and the connections between them weaken, reducing joint activity. This mechanism of neural pruning has been called anti-hebbian neural pruning (Földiák, 1990). Figure 6A illustrates our hypothesis and Figures 6B,C&D show experimental proof for this phenomenon. The mouth-head connection starts off weaker than the head-mouth connection. This is because the onset of mouth movements is later than that of head movements; when the head is moving the mouth is not moving but when the mouth moves the head can move in tandem. The strength of connections remains the same until the amount of mouth movement exceeds the amount of head movement. Once this point is reached, the mouth moves but the head is not moving as much; the mouth-head connection is weakened by anti-hebbian mechanisms but the head-mouth connection remains the same, albeit feeble. When mouth movements start declining the already weak connections weaken further. This is because movements of both types are low and at this point neither movement can be reliably accompanied by the other. This mechanism is similar to the mechanism effecting self-organization in the simulations of Pfeifer and colleagues (Marques et al., 2013).

In which part of the nervous system might we find these networks? Most studies on fetal motor development have focused on limb movements. Experiments on rat fetuses have shown that the spinal cord alone is sufficient to organize many aspects of spontaneous limb activity (Robinson et al., 2000; Kleven et al., 2004). This cannot be the case for mouth and head movements; instead, we propose the reticular formation region of the brainstem — a region thought to control early fetal movements (Shuleikina and Gladkovich, 1965; Gladkovich et al., 1980). In adult animals, the central pattern generator for rhythmical mouth movements is located in the reticular formation (Chandler and Tal, 1986; Nozaki et al., 1986; Hashimoto et al., 1989) and, this area consists of premotor interneurons that are directly controlled by the mouth area of the motor cortex (Nozaki et al., 1986). Fetal mouth movements could be under the control of

this region even if connections to the motor cortex have not formed. The reticular formation is also involved in adult head movement (Sirkin et al., 1980). Experiments on cat fetuses have shown that the reticular formation has one of the earliest appearances of electrical activity and controls early fetal movement (Gladkovich et al., 1980). Morphological comparison of fetal and infant neurons in this region shows that one major neural subtype, the multipolar giant cells, has dendritic structures with rich branching in the fetus (45-55 gestational days, 10-20 days from birth) and comparatively less branching in kittens 1-5 days and 30 days old (there was no difference in neural structure between infant age groups) (Gladkovich et al., 1980). Pruning of these branches could have occurred due to synaptic weakening by anti-hebbian mechanisms (Cheung et al., 2014). Further, the multipolar giant cells are specifically associated with motor pathways that control the face, head and neck muscles. This is the case in both fetus and infant (Gladkovich et al., 1980). In summary, the reticular formation region of the brainstem is associated with mouth and head movements in adult animals, is active early in fetal life and has a subpopulation of cells showing neural organization (synaptic pruning) from fetal to infant life.

The increase in neural organization with age is reflected in the behavior of the fetus. We have used entropy measures to quantify this progression (Figure 7). Early in development mouth and head motor behavior is unpredictable, movements occur in one of 5 ways – independent mouth and head movements, mouth movements that have an overlapping head movement, head movements that have an overlapping mouth movement and synchronous movements of the two types (Figure 7A). At this point, the system has high entropy. With age the number of states within a single session and therefore entropy, decreases (Figure 7A, B). Additionally, the entropy measure on each successive day is increasingly different from the entropy measure on the first day of testing (Figure 7C). The decrease in entropy with each successive day is due to the gradual elimination of overlap as mouth and head mature into decoupled motor modules. At the very end of gestation, with the exception of one or two stragglers, only independent mouth and head movements happen. This is observed in all pregnancies (Figure 7A&E).

Figure 8 summarizes our conceptualization of the differentiation of mouth and head motor modules. The environment of a fetus is the uterus and the fetus actively samples this environment through motor activity. Fetal experience is dictated by features of the uterine environment (e.g. space constraint and physical contact with siblings inhabiting the same space) and by features of the fetus itself (size of fetus, body shape of fetus, muscle connectivity, tactile receptor distribution, neural circuits of the fetal brain). Fetal behavior self-organizes through dynamic and reciprocal interactions between the brain, body and environment. In our example, we contend that inhibition from the developing brain, differential tactile receptor distribution on the fetal body and increasing space constraint from the uterine wall all serve to organize the connections between mouth and head movements, via anti-hebbian rules of neural pruning.

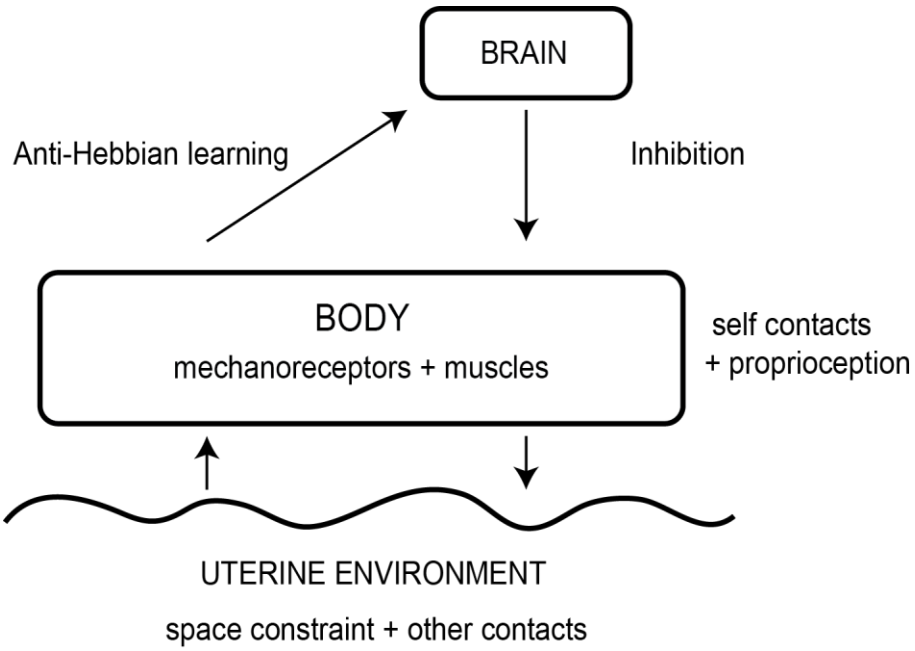


Figure 8. Proposed framework of mouth and head decoupling. Mouth and head modules self-organize through dynamic interactions between the body (musculoskeletal structure and mechanoreceptor distribution), brain (establishment of inhibitory mechanisms in the developing brain) and environment (increasing space constraint and tactile stimulation from siblings sharing the uterine space).

Methods

Subjects

We tracked four pregnancies and a total of nine fetuses. This included one pregnancy of a single baby, two pregnancies of twins, and one pregnancy of quadruplets. Of the nine fetuses, seven survived (two of the quadruplets were stillborn). There were no obvious differences in the sizes of the singleton, twins, or quadruplets at birth (including the stillborn infants). Our veterinary doctor performed necropsies on the stillborn animals and concluded that both were well developed and looked to be at term.

The four studied pregnancies were of two adult marmosets - three pregnancies of one female and one pregnancy of the other. Both animals were part of marmoset social groups housed at the Developmental Neuromechanics and Communication Laboratory, Princeton University. The two pregnant females came from different social groups and lineages. Animals were fed once daily with standard commercial chow (Harlan Marmoset Diet #TD.08482.PWD) supplemented with fresh fruits and vegetables. Additional treats (typically peanuts, cereal, dried fruit, raisins and marshmallows) were given during experimental sessions. The animals had ad libitum access to water. The colony room was maintained at a temperature of approximately 27°C and 50-60% relative humidity, with a 12L:12D light cycle. Previous to conducting the experiment, all animals were trained to leave their home cage in a transport box and familiarized with the testing room and testing equipment. All experimental sessions were in daylight hours - between 1400 and 1800. All procedures were in compliance with guidelines of the Princeton University Institutional Animal Care and Use Committee.

Ultrasonography technique

Ultrasonography tests were performed using a GE Voluson i Ultrasound machine. The testing method used was a modified version of the procedure developed by S.D. Tardif and colleagues (Jaquish et al., 1995). Each examination was conducted by two experimenters — one gently restrained the animal while the other carried out the ultrasound procedure. The animals were not anesthetized for the examination. Instead,

they were trained to accept gentle restraint and were rewarded with treats at the end of the testing period. We routinely tested all our adult females for pregnancy. Early pregnancy was detected by assessing the morphology of the uterus using the method described by Jaquish and colleagues (Jaquish et al., 1995). Once pregnancy was confirmed, ultrasounds were conducted every two weeks to monitor the development of the fetus. When the fetal skull became clearly visible, gestational age could be estimated by measuring the biparietal diameter of the skull (Jaquish et al., 1995). At this point, we started monitoring the fetus 2-3 times a week and as soon as the fetal face was clearly visible (usually around 95 days gestational age), examinations were recorded. A total of 53 sessions were recorded – 17 sessions for the singleton and 12 sessions for each of the other pregnancies. The examination was terminated if the animal showed significant resistance. A typical experiment ranged from 15-40 minutes. During these examinations the probe was manipulated to ensure that a frontal view of the face was attained at all times. In cases where there was more than one fetus, two fetuses were selected at random and each fetus was observed for 15-20 minutes. The ultrasound videos were captured at a frame rate of 29.97 and written on to DVD for later analysis. All exams were performed between 14:00 and 18:00 hours and the procedure was repeated until birth (usually around 146 days gestational age).

Identifying mouth and head movements

Video analysis was done using the Adobe Premiere Pro software, which enabled us to screen the videos frame-by-frame. Since the animals were not anesthetized for the ultrasound procedure, both the mother and the fetus could move, causing us to lose sight of the fetal face from time to time. Therefore, we first identified segments of the video during which the fetal face was clearly visible. Within these segments, we scored for mouth and head movements. A movement counted as a mouth movement when there was a clear separation of the upper and lower jaws (Figure 1B image panel, in red). The first video frame where the jaws separated counted as the beginning of mouth movement and the first video frame where the jaws came back together was the end point of the movement. If we lost focus of the face before the jaws came back together, the last frame of observation counted as the end of the movement. We did not separate between

different types of mouth movements. Individual mouth movements were considered to be part of the same movement unit if they were separated by 500ms or less (~15 frames). The 500ms criterion was justified by the bimodal structure of the inter-syllable interval distribution of the vocal output of marmoset neonates. In their calls the 500ms threshold separates the first mode of the distribution (representing the interval between syllables within a single call) from the second mode that represents the interval between the offset of the last and onset of the first syllables between two calls. Similar to mouth movements, the first video frame where the head moved away from baseline position was the beginning of movement and the first video frame where the head came back to baseline was the end point of movement (Figure 1B image panel, in blue — the lower line shows the baseline, the upper line marks the position of the head after movement). If we lost focus of the fetus before the head came back to baseline position, the last frame of observation counted as the end of the movement. We applied the same criterion used for the mouth when combining individual head movements into a single unit. Henceforth, when we refer to mouth or head movements we are speaking of movement units.

Exemplars

For the movement exemplars in Figure (1B), the image panel was generated by first using a custom-made MATLAB® program to split a chosen video clip into its component frames and then selecting those frames which best exemplified the movement. The movement timeline plot was generated using onset and offset information of the movements (method described above) and plotted using MATLAB®. The amplitude of the movement did not have any bearing on the analysis.

Quantifying mouth and head movements

The numbers of mouth and head movements were counted for each session. Since the amount of time the face was visible differed between sessions, we calculated the rate of mouth and head movements per second. Movement rates were compiled across all pregnancies and polynomial curves (one for mouth movements, one for head movements) were fitted to look at the trends across gestational time. To fit the curves, we first found the optimal degrees for polynomial fitting according to Akaike's information criterion

(Akaike, 1981a) (MATLAB® polydeg). These degrees were then used in a polynomial curve fitting function (MATLAB® polyfit) to generate optimal fits for mouth and head movement rates. The t-statistics and p-values for the trends were calculated using MATLAB® curve fitting toolbox. The p-value of the fit coefficients were also generated by the polyfit function. To check if the observed trends held at the subject level, we split the mouth and head movement rates by pregnancy and used the optimal polynomial degrees calculated with the entire dataset to generate polynomial curves. The curves were generated and diagnostic statistics carried out using the procedure described above.

Refuting the space constraint theory

The biparietal distance of fetal skull was measured using the GE Voluson i Ultrasound machine. A similar method has previously been used by Tardiff and colleagues (Jaquish et al., 1995). For the pregnancies with multiple fetuses, the average biparietal distance of the two infants imaged in that session was calculated. The fitted curves were generated using the same method used in previous sections. The biparietal distances of the same fetuses on day 1 of postnatal life were measured using a piece of string and a ruler (including those of the two stillborn fetuses from the quadruplet pregnancy). The individual curves for mouth movement rates, head movement rates and overlap rates are a compilation of the curves generated for individual pregnancies in previous sections (Figures 2D and 3C).

Overlap of mouth and head movements (all instances)

We counted the number of instances where mouth and head movements occurred together and divided this number by the total number of mouth and head movements. In this case, we included all instances of overlap. The overlap calculation included instances where mouth movements were followed by an overlapping head movement, instances where head movements were followed by an overlapping mouth movement and instances where mouth and head movements started simultaneously. The median percentage of overlap for each session was calculated and a polynomial curve was fitted to look at the trend across gestational time. To fit the curve, we first found the optimal degrees for polynomial fitting according to Akaike's information criterion (Akaike, 1981a)

(MATLAB® polydeg). This degree was then used in a polynomial curve fitting function (MATLAB® polyfit) to generate the optimal fit. The fit coefficient was also generated by the polyfit function. To check if the observed trends held at the subject level, we split the overlap rates by pregnancy and used the optimal polynomial degrees calculated with the entire dataset to generate polynomial curves. The curves were generated and diagnostic statistics carried out using the procedure described above.

Permutation Test

The total number of mouth and head movements was seen to decrease through gestation. To confirm that the observed decline in mouth and head movement overlap was independent of the general decline in movement, a permutation test was performed. We took each session and independently reshuffled both the head movements and mouth movements, keeping the duration and latency distributions intact. We then independently resampled with replacement, keeping the number of durations and intervals the same as the original dataset. Overlap between mouth and head movements (all instances) was then calculated. This was done for every session and average percentage of overlap was calculated for each gestational age (same procedure as that used in our original calculation). This procedure was repeated 1000 times and the upper 97.5 and lower 2.5 percentiles were computed to generate the confidence interval. We expected this analysis to show the pattern of overlapping movements across gestation that would have resulted due to chance (rather than due to mouth and head movements being linked).

Duration of mouth and head movements

Mouth and head movement durations were calculated by subtracting a movement's offset time from its onset time. Median mouth and head movement durations for each session were calculated and compiled across all sessions, for all pregnancies. To test if there was a change in the duration of mouth and head movement through gestation we used the curve fitting method used for movement and overlap rates.

Overlap of mouth and head movements (directionality)

Instances of overlap were split based on the order of the movements. For every session,

the number of mouth-followed-by-head overlaps was divided by total number of mouth movements; the number of head-followed-by-mouth overlaps was divided by the total number of head movements. Both curves were fitted using the curve fitting method used for movement and overlap rates. The 95% confidence interval of the fit was calculated using MATLAB® polyconf function. The permutation test was done in the way described previously. The one difference being, overlap was calculated using the method described in this section. To check if the observed trends held at the subject level, we split the overlap rates by pregnancy. Here too, the curves were fitted using the curve fitting method used for movement and overlap rates. The shaded regions are the 95% confidence intervals of the fits.

Information theory Analysis

Each movement we assigned to one of 5 states: independent mouth movement (State1), independent head movement (State2), mouth movement followed by an overlapping head movement (State3), head movement followed by an overlapping mouth movement (State4), and mouth-head movements with synchronous onset (State5).

The percentage of each movement state in one session was calculated by dividing the total number of movements of a particular state by total movement in the session (e.g. number of state1 movements in session 1/ total number of movements in session 1). This was repeated for all sessions.

Shannon Entropy (H) for each session was calculated using the following formula:
$H(X) = -\Sigma_i P(x_i)\log_2 P(x_i)$ where i=1 to 5 (corresponding to the 5 states). To test if there was a change in entropy through gestation we used the curve fitting method used for movement and overlap rates.

Kullback-Leibler Divergence (D_{KL}) of *Q* from *P* was calculated using the following formula: $D_{KL}(P||Q) = \Sigma_i P(i)\log_2 P(i)/Q(i)$ where i=1 to 5 (corresponding to the 5 states); *P* was the state distribution for the first testing session (gestational day 93); *Q* was the state distribution for every testing session (gestational days 93, 95….147).

Calculations made for the state space diagrams: The gestational time after which the fetal face was clearly visible was split into four gestational periods. We determined the number of days in each period (for every pregnancy) by counting the days between the first day the fetal mouth was clearly visible and the gestational age at birth and then dividing the number of days in four. The average length of a gestational period, across all pregnancies, was 12 days (± 2 days). For each pregnancy, from each quadrant, the most common state distribution pattern was identified and a session having that pattern was chosen. The one-step transition probabilities were then calculated using the following formula: $P_{kj} = n_{kj}/\Sigma_i n_{ki}$ where i=1 to 5.

Chapter 2

The prenatal origins of “innate” vocalizations in marmoset monkeys

Summary

In primates, a number of vocal behaviors are present at birth. These behaviors are often dubbed as ‘innate’ but how they develop remains mysterious. As opposed to simply appearing de novo at birth, we propose that vocalizations begin to self-organize prenatally through the interactions between spontaneous neural activity, fetal biomechanics and the uterine environment. We investigated how fetal movements relate to the postnatal mouth movements required for vocal production by performing ultrasound imaging on awake, pregnant marmoset monkeys (*Callithrix jacchus*). We found that signature features of marmoset infant calls emerge prenatally as distinct patterns of orofacial movements: Late in gestation, the duration and syllable number of the phee call and twitter call — calls produced by marmoset monkeys on the first postnatal day — are evident in the orofacial movements of fetal marmosets. Our study shows that aspects of vocal behaviors in marmosets have a period of prenatal development.

Introduction

Behaviors present at birth are often dubbed ‘innate’ or ‘inherited’ (Lorenz, 1950; Arbib et al., 2008; Hammerschmidt and Fischer, 2008). On closer inspection, as opposed to simply appearing de novo at birth, these same behaviors are seen to have a long period of in-utero development. For instance, the universality of human facial expressions has led to claims that they are ‘innate’ (Izard, 1994; Matsumoto and Willingham, 2009). However, facial expressions resembling infant smiling, scowling, pouting, hiccupping and yawning are frequent at the end of pregnancy (Kurjak et al., 2003). Lincoln and colleagues have examined the prenatal development of cry-like and laughter-like

expressions (Reissland et al., 2011). The facial expressions of infants have been broken down into constituent facial muscle movement combinations or action units. Their paper used these action units to study fetal facial gestalts. They saw a change from a few action units observed in isolation in the second trimester, to recognizable cry and laughter expressions in the third trimester of gestation. In other animals too, the myth of 'innateness' has been torn down after inspection of the fetal stage. Pecking behavior in chicks, once thought to be 'instinctive' is now known to develop in the egg, through interactions between spontaneous neural activity, fetal biomechanics and the uterine environment. Perturbations to any of these elements can change the course of development (Kuo, 1932a, c; Kuo, 1932b; Lehrman, 1953).

In primates, a number of vocal behaviors are present at birth. Crying sounds, for instance, are ubiquitous among newborn primates; separation-induced cries of humans and other primates have broadly similar acoustic structure and are thought to be essentially the same sounds (Newman, 1985). These behaviors, and other vocal behaviors present at birth, are also often dubbed 'innate' (Arbib et al., 2008; Hammerschmidt and Fischer, 2008). We suggest that vocalizations present at birth develop in the months of fetal life. To show that this is case, what is needed, is a longitudinal study using a non-invasive approach to measure discrete fetal vocalization-related movements.

In both human and nonhuman primates, sound is produced in a similar manner: Diaphragm action pushes air from the lungs and causes vibration of the vocal folds of the larynx. This sound source travels up the vocal tract, where the movements of face effectors (jaws, lips, tongue) act as a filter, passing acoustic energy at some frequencies and attenuating energy at other frequencies, according to their size and shape. The different mouth movements, along with changes in vocal fold tension and respiratory power, are what give rise to different sounding vocalizations (Ghazanfar and Rendall, 2008; Ghazanfar and Takahashi, 2014). Air is required for sound production and under normal conditions fetuses do not produce sounds (Ryder, 1943). Nonetheless, for vocalizations to be produced at birth, some actions of the vocal apparatus must be developing in-utero. Not all parts of the fetal vocal apparatus can be observed non-

invasively but fetal mouth (jaw) movements can be observed. We have used ultrasound imaging on awake, pregnant marmoset monkeys (*Callithrix jacchus*), to track fetal movements of the mouth, from the day the mouth is first visible to the day before birth. The approach has allowed us to observe not just the start and end points of motor development, but the entire trajectory in individual pregnancies, in a non-invasive manner. We then compared fetal mouth movements to the signatures of infant vocalizations on the first day of birth.

Marmoset monkeys are an ideal model system for this investigation because they have roughly the same body plan as humans and have homologous brain structures (Burman and Rosa, 2009; de la Mothe et al., 2012). Simulations have highlighted the importance of body morphology for fetal motor development (Mori and Kuniyoshi, 2010; Mori, 2012; Yamada and Kuniyoshi, 2012). Given the similarity of primate bodies, it is likely that the basic developmental trajectory would be characteristic of human fetal development as well. These monkeys also breed readily in captivity and regular ultrasound imaging can be done on them with no adverse effects on the fetus (Jaquish et al., 1995).

Marmosets and most other nonhuman primates are born precocial and develop at a much faster rate than humans. In addition to cries, just born marmosets produce calls that are structurally comparable to adult calls (Bezerra and Souto, 2008); these calls can be classified based on their distinct sound, spectral features, temporal features, durations and syllable numbers. The most prolific are twitters, trills, phees and in case of the infants, cries. Adult animals produce twitter and trill calls when in visual contact with conspecifics; phees are used to maintain auditory contact with out-of-sight conspecifics. In fact, adult marmosets phee call in bouts of contingent exchanges and their turn-taking dynamics are similar to those of human conversations (Takahashi et al., 2013). Infants develop adult structure and correct use of this call within the first two months of postnatal life (Takahashi et al., in press). Very young infants produce a large number of cries but a few weeks after birth, they start to produce mostly phee calls. An infant development study from our group shows that the cries are immature forms of phee calls (Takahashi et

al., in press). The study also shows that the transformation of cries to phee calls cannot be solely attributed to growth; contingent vocal interactions from parents influence the timing of transition. Thus, as in humans early infant vocalizations lay the foundation for more mature vocalizations (and this transformation is influenced by social feedback).

In an equivalent manner, aspects of fetal behavior must lay the foundation for infant vocal behavior. The main prediction of our study was that signature features of marmoset infant calls — features such as duration and syllable number — emerge prenatally as distinct patterns of orofacial movements. Figure 9 provides exemplars to show that this is indeed the case. Figure 9A depicts a typical marmoset phee call (in red) and twitter call (in blue). Each strip of images comes from an exemplar call. The tracings have been used to indicate the duration and syllable number of the calls. Figure 9B shows fetal mouth movements that have duration and syllable number features of the two call types (phee-like movement in blue, twitter-like movement in red).

Results

To understand the developmental continuity of vocal behavior from fetal to infant stages, we first characterized the vocal behavior of infant marmosets. Our focus was on call production. Infant vocalizations have characteristic spectral prints, durations and syllable numbers. Figure 10 shows the four most common types of infant calls (Bezerra and Souto, 2008) — phee (Figure 10A); twitter (Figure 10B); trill (Figure 10C); cry (Figure 10D). The top panel of each section has an image strip from the selected exemplar; the bottom left panel shows the spectrogram of the same call and the bottom right figure has a tracing of the orofacial movements accompanying the call. Under normal conditions an animal does not vocalize in-utero (Ryder, 1943). Therefore, we cannot use the acoustic features of a call to identify pre-cursors of vocalizations in the fetal stage — we defined the calls based on their duration range and syllable number range, and used these measures to sieve fetal movements.

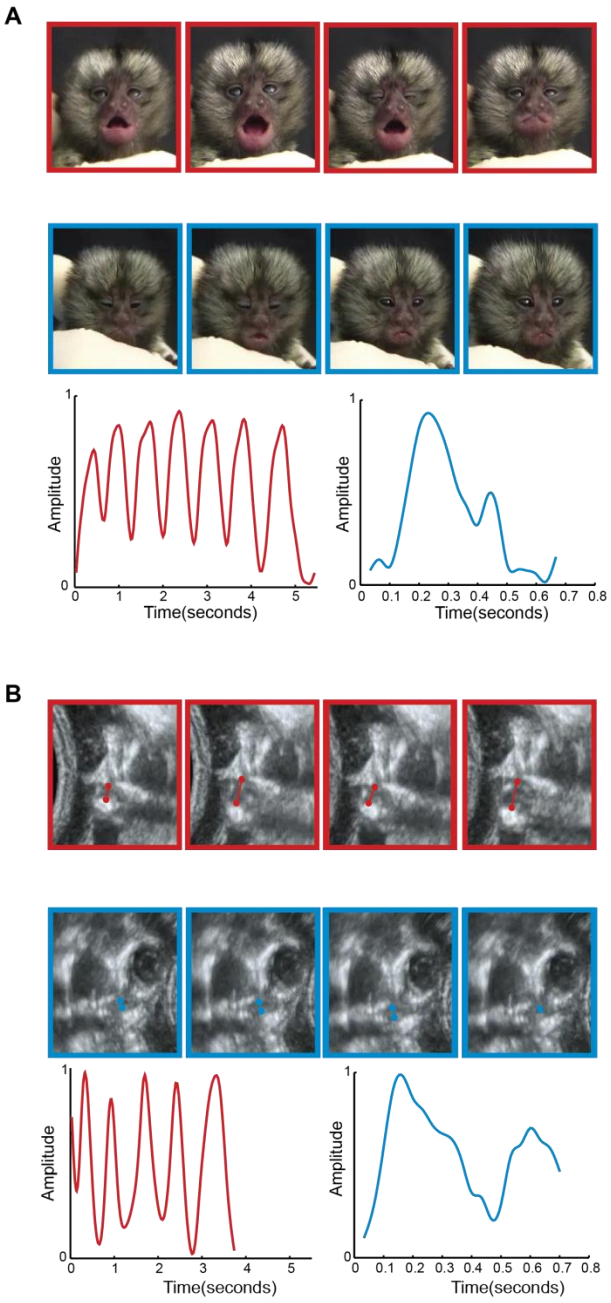


Figure 9. Exemplars to illustrate that infant call characteristic can be seen in fetal movements. The top panel in each set shows still frames extracted from video clips. The bottom panel of each set has tracings of the mouth movements. The tracings have been used to indicate duration and syllable number. The x-axis has the duration information. The amplitude of movement had no bearing in our analysis so the y-axis value provides no additional information. (A) Depicts a typical marmoset infant phee call (in red) and twitter call (in blue). (B) Shows fetal mouth movements that have duration and syllable similar to that of phee call (in red) and twitter call (in blue).

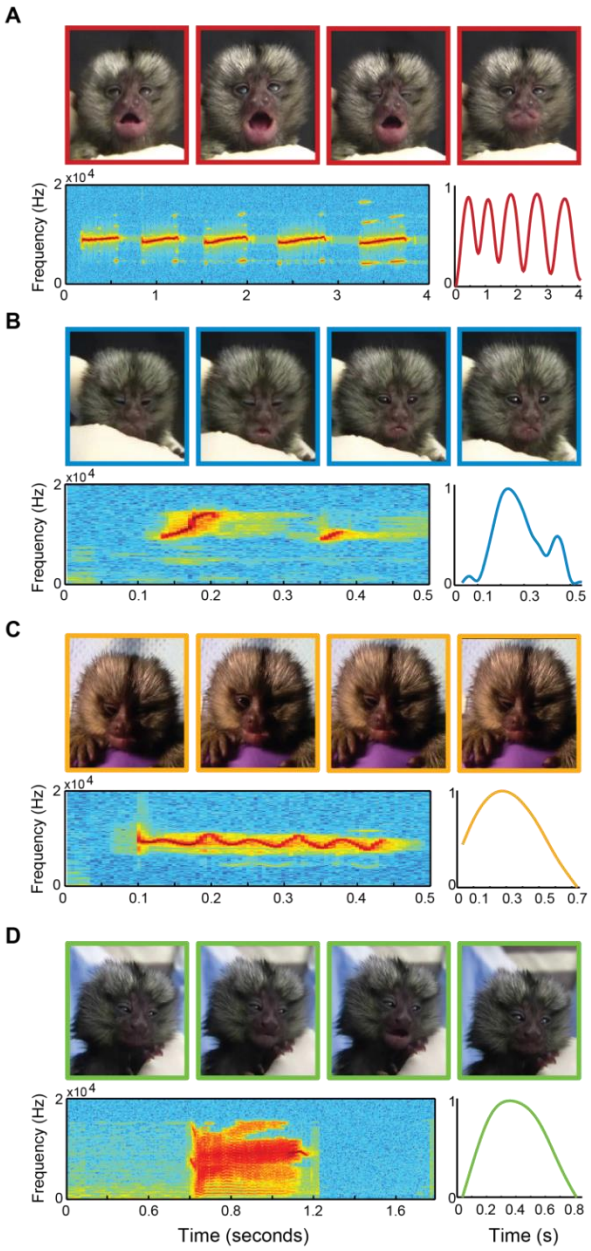


Figure 10. Four common types of vocal signals produced by newborn marmoset infants. The top panel of each section has an image strip from the selected exemplar; the bottom left panel shows the spectrogram of the same call and the bottom right figure has a tracing of the mouth movements accompanying the call. Note that there is discrepancy in the duration information acquired from the audio (spectrogram) and video (tracing) (A) Phee call: audio duration 3.58s; video duration 3.94s; syllable number 5 (B) Twitter call: audio duration 0.25s; video duration 30.43s; syllable number 2 (C) Trill call: audio duration .53s; video duration .68s; syllable number 1 (D) Cry: audio duration .58s; video duration .80s; syllable number 1.

A total of seven infants were tested (4 males, 3 females). Five of the infants had the same parents and two of them another set of parents. Infant testing was started ~24 hours after birth and audio plus video measures of their vocalizations were made. For the audio recordings the infant was placed within a small metal cage, on a flat piece of foam. The experimenter then left the room and all vocalizations produced by the infant were recorded for a period of 5 minutes. After, the infant was removed from the cage, held by one experimenter, and a second experimenter took video recordings of the infant face. These recordings lasted for 5-10 minutes. The infant typically produced different types of vocalizations during this period. The procedures were repeated every day for the first week after infant birth.

To analyze the audio recordings a custom made MATLAB routine was used to determine the onset and offset timings of each call syllable. Multiple syllables were combined into the same call if they had the same spectral prints and if they were separated by less than 500ms (DiMattina and Wang, 2006). Durations and syllable numbers of all calls of the same type were then compiled in order to determine the duration and syllable number profile of each call type.

We took the same measures for the video recordings. A frame-by-frame analysis was done for the video recordings using Adobe Premiere. The first frame of mouth opening counted as the onset of movement and the first frame of mouth closure counted as the offset of movement. To identify the type of call, the audio signal was extracted from the video recording and the spectrogram of the call was examined using Adobe Audition. By comparing spectral features of the audio signal accompanying the video, orofacial movement signatures of each call type could be identified. Multiple movements were treated as part of the same call if they had the same orofacial signature, the same spectral features and were separated by less than 500ms (same as with audio recordings alone). The frame rate of the video recording was 29.97 frames per second; therefore, movements separated by less than 15 frames were classified as part of the same call. Individual calls were clipped out of the length of recording and the duration time and syllable number of each call was recorded. The number of syllables in a call counted as

the number of individual mouth movements in the call. Durations and syllable numbers of all calls of the same type were then complied in order to determine the duration and syllable number profile of each call type.

For our analysis, we used the first day of audio recording and the entire first week of video recording. We found that marmoset infants fall asleep a lot on day 1 (just like any other newborn), especially when held. In order to get a large enough sample size, we had to include the entire first week of video recording. Both our audio and video recordings consisted of infant phees, twitters, trills and cries. The audio recordings on day one consisted of a total of 603 calls (approximately 20% phees, 18% twitters, 13% trills and 41.96% cries). The video recordings consisted of a total of 126 calls (approximately 33% phees, 21% twitters, 22% trills and 25% cries), 20 of these calls were from postnatal day one.

In order to use the infant audio measures to structure our search for vocalization pre-cursors in the fetal stage, we had to first test if the measures taken from the audio signal could be equated with the measures taken from the recorded orofacial movements (fetal measurements were taken from ultrasonography videos). We could have just used measures taken from the video to define our call parameters but in terms of experimental set-up, the process of recording audio signals is more comparable to fetal measures than the process of recording the videos. For the audio, the infant was placed on the foam and allowed to move around and vocalize on it's own volition. Whereas for the videos, the infant was held by an experimenter and sometimes coaxed into vocalizing by gentle stroking. Further, we were most interested in call signatures on the first day of postnatal life. As stated earlier, the video recordings from the first day would not have given us sufficient calls to define our parameters in a satisfactory manner. Also, the number of recorded audio calls being more numerous, they are probably more representative of infant call characteristics.

To compare the two signals, we extracted audio from the video recordings and measured call duration and syllable number using the same method used for the audio-alone

recordings. For each call, we looked at the difference between the extracted audio signal and the mouth movements in the video. In case of duration, the video measures were consistently longer than the audio. The exemplars in Figure 10 illustrate this — compare the spectrogram on the left to the mouth movement tracing on the right [(A) Phee call: audio duration 3.58s; video duration 3.94s (B) Twitter call: audio duration 0.25s; video duration 30.43s (C) Trill call: audio duration .53s; video duration .68s (D) Cry: audio duration .58s; video duration .80s]. To compensate for the duration difference, we calculated the difference between the video and audio duration and grouped the differences by call type. We calculated the median duration discrepancy for the phee and the 90 percentile discrepancies for the other call types. The median value was added to the audio duration of all phee calls and the 90 percentile values were added for the other call types. The number of syllables remained the same across the audio and video measures; no adjustment was required.

We then set out to define the infant P1 calls using duration and syllable number. Figure 11A shows the probability density estimates of call durations, after discrepancy adjustment. Phee calls are in red, twitter in blue, trill in yellow and cry in green. The black bars mark median durations of each call. The minimum duration of any call type is ~500ms and the maximum duration of any type of call is ~6.5 seconds. The phee call distribution has a median value of 3.69 seconds and spans from around 500ms to 6.5 seconds [SD: 1.49]. The median durations of the other calls are more closely clustered [twitter — median: 0.96, SD: 0.43; trill — median: 0.69, SD: 0.16; cry — median: 1.17, SD 0.96]. The heat maps in Figure 3B show how the calls cluster based on duration. In the heat maps, the color red shows the durations values with highest call density and blue, the region of least density. Phee calls tend to be of long duration compared to the other three calls. Comparing the distributions, it is apparent that infant phee call is most easily distinguishable.

Figure 11C shows the probability density estimates of the call syllables. Here too, phee calls are in red, twitter in blue, trill in yellow and cry in green; the black bars mark

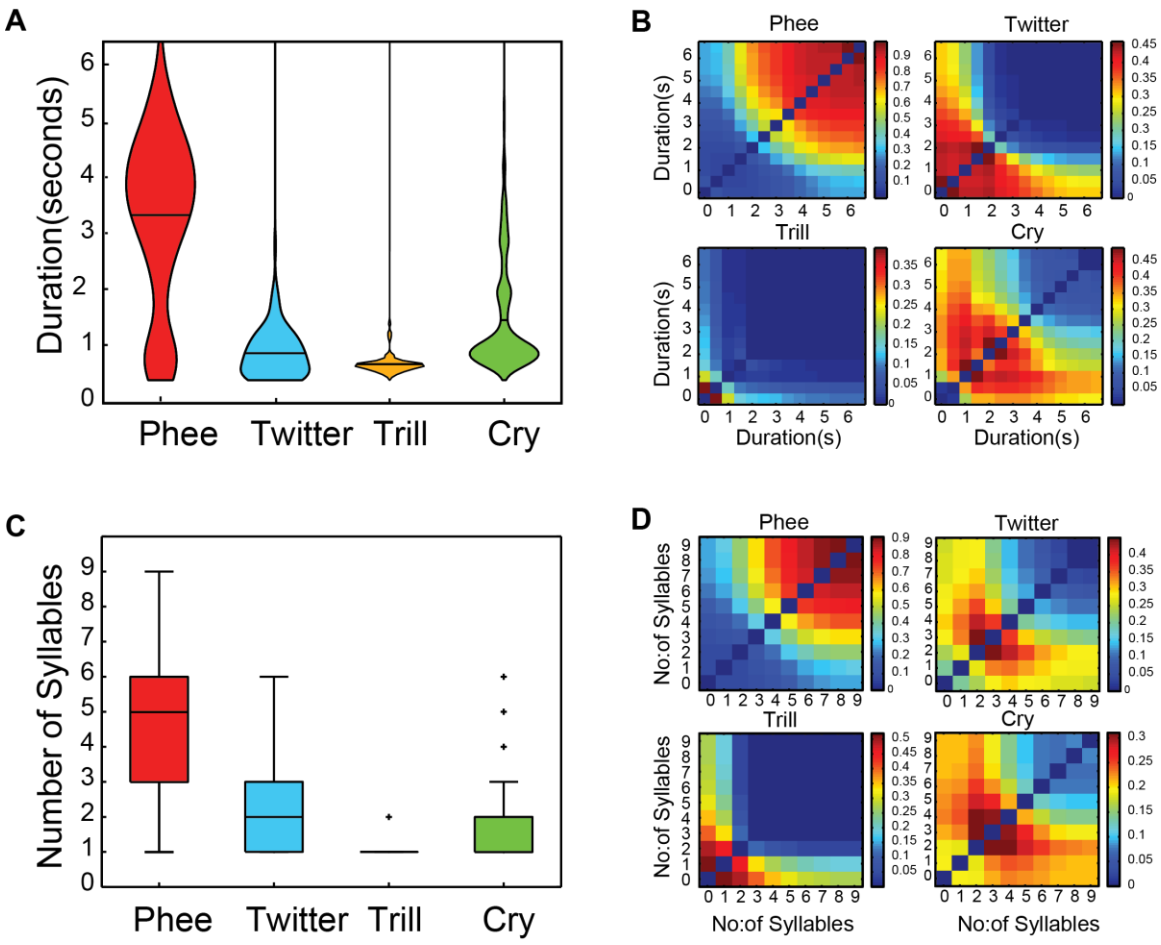


Figure 11. Duration and syllable number distributions of infant call types. (A) Probability density estimates of call durations, after discrepancy adjustment. Phee calls are in red, twitter in blue, trill in yellow and cry in green. The y-axis shows duration in seconds. The black bars mark median durations of each call. (B) Heat maps to show how well the calls can be discriminated by duration information alone. The color red indicates duration values with highest call density and blue, the region of least density. Note that phee is the most distinct call. (C) Probability density estimates of call syllable numbers. Phee calls are in red, twitter in blue, trill in yellow and cry in green. The y-axis shows syllable number. The black bars mark median number of syllables of each call. (D) Heat maps to show how well the calls can be discriminated by syllable information alone. The color red indicates syllable number values with highest call density and blue, the region of least density. Note that here too, phee is the most distinct call.

median syllable number. The minimum syllable number of any call type is 1 and the maximum syllable number of any type of call is 9. The phee call syllable distribution has a median value of 5 and spans from 1 to 9 syllables [SD: 2.01]. As with duration, the median syllable numbers of the other calls are more closely clustered [twitter — median: 2, SD: 1.28; trill — median: 1, SD: 0.22; cry — median: 1, SD 0.99]. The heat maps in Figure 11D show how the calls cluster based on syllable number. In the heat maps, the color red shows the syllable values with highest call density and blue, the region of least density. Phee calls have more syllable numbers compared to the other three calls. Comparing the distributions, it is apparent the infant phee call is most easily distinguishable based on syllable number (same as duration).

To be able to look at the fetus and match the duration and number of mouth movements (equivalent of syllable number) to a specific infant call, we had to come up with criteria for defining each infant call type. Following is the method we used for setting the parameters (1) Not knowing what numbers of call precursors are present in-utero, we ignored the discrepancy in numbers of infant call types and assumed that the quantities of all call types are equal (2) For every duration (500ms - 6.5s) or syllable number (1-9), we calculated the likelihood that a call was of a particular type [For instance, if the duration is 6s the likelihood that the call is a phee call is 100%] (3) We set the parameters for each call to ensure maximum possible separation from other call types. For the ranges we settled on, the median duration or syllable number of a call type served as the lower bound; the upper bound varied depending on the feature distribution (Figure 11) and the best fit based on the above criteria. Table 1 shows the duration bounds we used for each call and Table 2 shows the syllable number bounds we used for each call type. For each range the probability that a call within that range is of a particular type has also been listed.

We decided to look only for phee and twitter markers when examining ultrasound movements. We picked the phee call, as it has distinct duration (Table 1) and syllable number (Table 2) [The probability that a call within the selected duration range is a phee call is 0.91; the probability that a call within the selected syllable range is a phee call is

	Phee Criteria 3.69 – 6.5s	**Twitter Criteria 0.96 – 1.5s**	**Trill Criteria 0.69 – 1s**	**Cry Criteria 1.17 – 3s**
Phee likelihood	**0.91**	0.06	0.04	0.20
Twitter likelihood	1.19e-06	**0.40**	0.27	0.34
Trill likelihood	0	0.06	**0.44**	0.05
Cry likelihood	0.09	0.48	0.24	**0.42**

Table 1. Duration criteria for defining infant call types. The column titles have the call type and duration criteria of that particular call type. The values listed in the table are the likelihoods that a call meeting the duration criteria is of a particular type. The highlighted values are the ones where the criteria and call type match (e.g. phee call duration criteria and the likelihood that a call within that range is a phee call).

	Phee Criteria 5 - 9	**Twitter Criteria 2 - 3**	**Trill Criteria 1**	**Cry Criteria 1 - 4**
Phee likelihood	**0.8**	0.13	0.08	0.14
Twitter likelihood	0.10	**0.45**	0.20	0.34
Trill likelihood	0	0.11	**0.52**	0.27
Cry likelihood	0.10	0.31	0.20	**0.24**

Table 2. Syllable number criteria for defining infant call types. The column titles have the call type and syllable number criteria of that particular call type. The values listed in the table are the likelihoods that a call meeting the syllable number criteria is of a particular type. The highlighted values are the ones where the criteria and call type match (e.g. phee call syllable number criteria and the likelihood that a call within that range is a phee call).

0.8]. Based on duration or syllable number alone, twitter calls are not as clearly distinguishable (Table 1, 2). However, to show that our observations are not restricted to phee calls alone, we also looked for twitter call precursors. We chose twitters, and not any of the other two calls, because they provide the most contrast from phee calls: (1) Twitter calls are short duration calls with multiples syllables while phee calls are long calls with multiple syllables (Figure 11) (2) Cry calls are thought to be immature phee calls (Takahashi et al., in press) (3) Trill calls are similar to single syllable phees (Figure 10A,C; Figure 11A) and they cannot be easily separated from the other calls (Table 1,2).

Our fetal experiments aimed to observe signature features of phee and twitter infant calls emerge prenatally as distinct patterns of orofacial movements. We monitored spontaneous mouth movements in nine fetuses of two pregnant marmoset monkeys using ultrasonography. The seven infants we tested were all observed as fetuses. Of the four pregnancies, one was of a singleton, two of twins and one of quadruplets. Of the nine fetuses, seven survived (two of the quadruplets were stillborn). There were no obvious differences in the sizes of the singleton, twins, or quadruplets at birth (including the stillborn infants). Our veterinary doctor performed necropsies on the stillborn animals and concluded that both were well developed and looked to be at term. Observations of fetal mouth movements started on the first day the face was clearly visible (~ gestational day 95) and lasted until the day before birth (~ gestational day 146). The analysis in this paper is from observations conducted (roughly) on every other day of gestation. In total, we had 64 sessions of testing (14 sessions from the first pregnancy; 17 sessions from the second; 17 sessions from the third; 16 sessions from the fourth pregnancy). In each session, the ultrasound probe was positioned in a manner that enabled us to see a frontal view of the fetal face, and movements were monitored for around 15 to 45 minutes (the ultrasound was terminated if the mother became intolerant of the procedure, hence the variability in session length). When there were multiple fetuses we picked two fetuses at random and tracked the mouth and head movements of each fetus for 15 to 20 minutes.

We then did a frame-by-frame analysis of the ultrasound videos using Adobe Premiere. The procedure used for analyzing the ultrasound videos was near identical to the method

employed for the infant videos. We identified sections of video where the mouth was visible and clipped out the segments of mouth movement. A movement counted as a mouth movement when there was a clear separation of the upper and lower jaws (Figure 9B image panel). The first video frame where the jaws separated counted as the beginning of mouth movement and the first video frame where the jaws came back together was the end point of the movement. If we lost focus of the face before the jaws came back together, the last frame of observation was marked as the end of movement. The amplitude of movements did not have a bearing in our analysis. Next we grouped individual movements into a single movement unit if they were separated by 500ms or less. The 500ms criterion was determined based on the vocal behavior of infant marmosets (DiMattina and Wang, 2006). The number of individual movements within a movement unit was the 'syllable number' of the unit. Henceforth, when we refer to mouth movements we are speaking of movement units and syllable numbers stand for the number of individual movements within a unit.

We were only interested in vocalization-related movements so we first filtered out all other types of fetal movement. This was done by only including fetal movements within the range of vocalization durations (upto 6.5 seconds) and vocalization syllable numbers (1-9 syllables). These movements totaled to 93.33% of all fetal movements. Our analysis from this point on, concentrates on this narrowed-down pool of fetal mouth movements.

Focusing on these 'vocalization-related' movements we first looked at the general change in duration and syllable number through gestation. Figure 12A shows the change in duration of these movements through gestation. Each black dot on the graph represents the median movement duration of one session; the black squares and black triangles show the upper 90 and lower 10 percentile points, respectively, for each session. The curves were generated by first calculating the optimal degree for polynomial fitting according to Akaike's information criterion (AIC) (Akaike, 1981b), for the median duration values. This polynomial fitting degree was used to generate the other two curves. [Using AIC, for

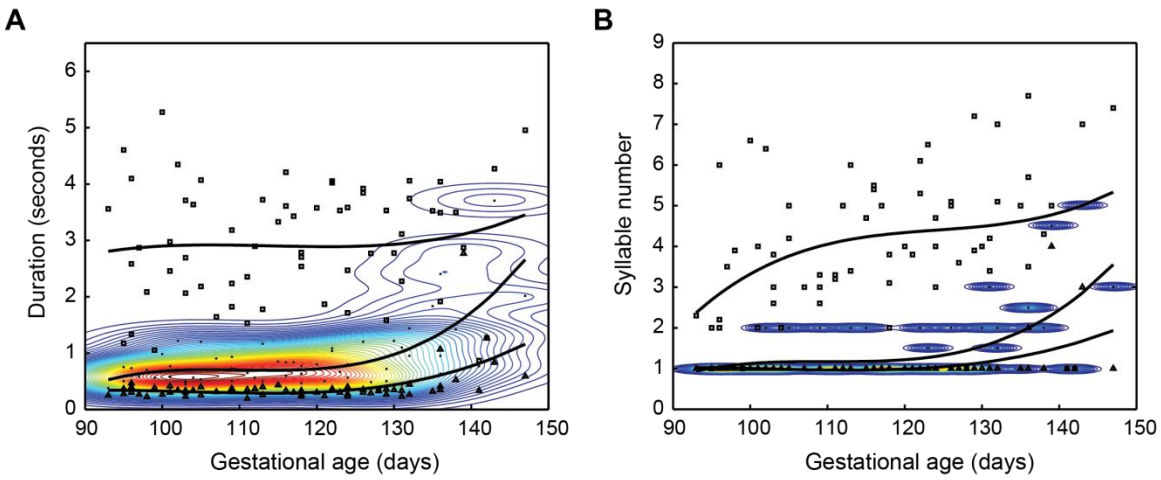


Figure 12. Trajectory of mouth movement duration and syllable number through gestation. (A) Duration of mouth movements through gestation. The x-axis represents gestational age in days. The y-axis represents duration in seconds. Each black dot on the graph is the median movement duration of one session; the black squares and black triangles show the upper 90 and lower 10 percentile points, respectively, for each session. The curves were generated by first calculating the optimal degree for polynomial fitting for the median values according to Akaike's information criterion (AIC) [a polynomial curve of degree 3 was found to be in the best fit]. The background contour map indicates durations with highest density of movements. The color red indicates syllable number values with highest call density and blue, the region of least density. Observe that the movement duration increases through gestation. (B) Syllable number of mouth movements through gestation. The x-axis represents gestational age in days. The y-axis represents syllable number. Each black dot on the graph is the median syllable number of one session; the black squares and black triangles show the upper 90 and lower 10 percentile points, respectively, for each session. The curves were generated by first calculating the optimal degree for polynomial fitting for the median values according to Akaike's information criterion (AIC) [a polynomial curve of degree 3 was found to be in the best fit]. The background contour map indicates syllable numbers with highest density of movements. The color red indicates syllable number values with highest call density and blue, the region of least density. Density estimates are not usually used for discrete data; we added the colored contours only to highlight the trend. Observe that the movement syllable number increases through gestation.

the median values a polynomial curve of degree 3 was found to be the best fit: second order coefficient: 4.09e-05; for the upper 90 values: second order coefficient: 1.77e-05; for the lower 10 percentile values: second order coefficient: 1.26e-05]. The colored contours show the density estimate for median duration, through time. The red coloring marks regions of highest density. As seen in this figure, the duration of mouth movements show an increase from the time of movement onset to the end of gestation.

Figure 12B shows the change in syllable number of 'vocalization-related' movements through gestation. Each black dot on the graph represents the median syllable number of one session; the black squares and black triangles show the upper 90 and lower 10 percentile points, respectively, for each session. The curves were generated by first calculating the optimal degree for polynomial fitting according to Akaike's information criterion (AIC) (Akaike, 1981b), for the median syllable number values. This polynomial fitting degree was used to generate the other two curves. [Using AIC, for the median values a polynomial curve of degree 3 was found to be the best fit: second order coefficient: 4.73e-05; for the upper 90 values: second order coefficient: 5.23e-05; for the lower 10 percentile values: second order coefficient: 1.23e-05]. Density estimates are not usually used for discrete data; we added the colored contours only to highlight the trend. As seen with duration, the syllable numbers of mouth movements show an increase from the time of onset to the end of gestation.

What does this trend mean for the development of specific call types? To know this we inspected the fetal movements for duration and syllable number signatures of phee calls and twitter calls. First we divided the tested period into four roughly equal quadrants — days 93-105 constitute the first gestational period, days 106-118 the second period, days 119-131 the third period and days 132-148 the fourth gestational period. We then compared the percentage of movements that matched the duration and syllable number characteristics of phee and twitter calls between quadrants.

Given that infant phee calls are long in length (Figure 11A, in red) and the durations of fetal movements increase with increasing age (Figure 12A), we expected to see a rise in

the number of fetal movements that match the infant phee duration profile (in the infants the likelihood that a call within our 3.69-6.5s duration criteria is a phee call is .91). In Figure 13A, each red bar represents the percentage of movements falling within the phee bounds, in a given gestational quadrant. The plotted errors are the standard errors of mean. As expected, the percentage of movements fitting the phee duration criteria increases through gestation [two-tailed Fisher's exact test: group comparison: $p = .04$; pairwise comparison quadrant 1 and 2: $p = .04$; quadrant 1 and 3: $p = .66$; quadrant 1 and 4: $p = .29$; quadrant 2 and 3: $p = .06$; quadrant 2 and 4: $p = .02$; quadrant 3 and 4: $p = .15$]. The increase in movements with long 'phee-like' durations could be a result of the strengthening of fetal musculature with increasing age, allowing the fetus to produce longer movements with time.

Twitter calls on the other hand are short duration calls (in the infants likelihood that a call within our 0.69-1.5s duration criteria is a twitter call is .40), what we see with fetal movements that match this call duration was unexpected. The percentage of movements that match the twitter duration profile also increases with increase in fetal age (Figure 13B). In Figure 13B, the blue bars represent the percentage of movements falling within the twitter bounds, in each quadrant. The plotted errors are the standard errors of mean [two-tailed Fisher's exact test: group comparison: $p = .007$; pairwise comparison: quadrant 1 and 2: $p = .78$; quadrant 1 and 3: $p = .05$; quadrant 1 and 4: $p = .03$; quadrant 2 and 3: $p = .008$; quadrant 2 and 4: $p = .01$; quadrant 3 and 4: $p = .34$]. The increase in movements with the shorter 'twitter-like' duration cannot be explained by a simple increase in strength of musculature; it is indication that what we are observing is the establishment of preferred mouth movement durations, which then dictate the length of infant calls.

For syllable number, we see a trend similar to that observed in durations. Matching the profile observed in Figure 13A, we see an increase in the percentage of movements that have syllables within the infant phee syllable bounds (in the infants the likelihood that a call within our 5-9 syllable criteria is a phee call is .80). In Figure 14A, each red bar represents the percentage of movements falling within the phee syllable bounds, per

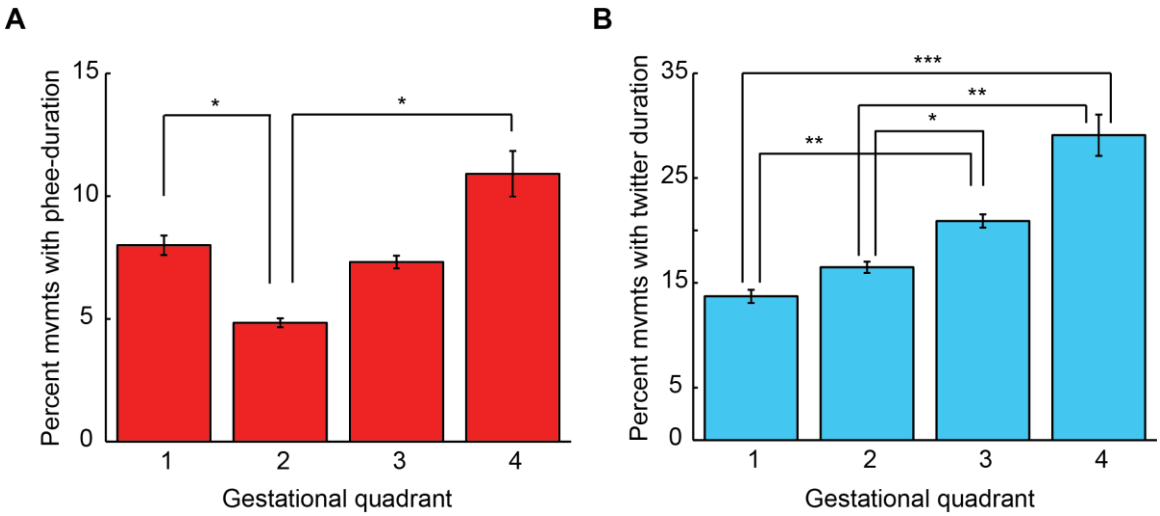


Figure 13. Percentage of fetal movements matching infant phee and twitter call duration signatures. The tested period into four roughly equal quadrants — days 93-105 constitute the first gestational quadrant, days 106-118 the second quadrant, days 119-131 the third quadrant and days 132-148 the fourth quadrant period. (A) Each red bar represents the percentage of movements falling within the phee bounds, in a given gestational quadrant. The plotted errors are the standard errors of mean. The significance test was done using Fisher's exact test. (B) Each blue bar represents the percentage of movements falling within the twitter bounds, in a given gestational quadrant. The plotted errors are the standard errors of mean. The significance test was done using a two-tailed Fisher's exact test.

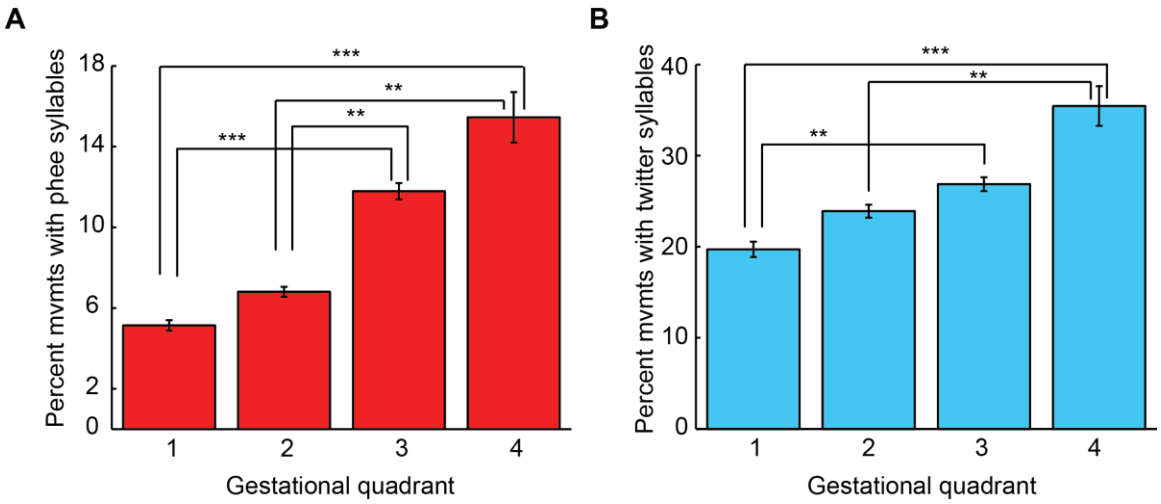


Figure 14. Percentage of fetal movements matching infant phee and twitter call syllable number signatures. The tested period into four roughly equal quadrants — days 93-105 constitute the first gestational quadrant, days 106-118 the second quadrant, days 119-131 the third quadrant and days 132-148 the fourth quadrant period. (A) Each red bar represents the percentage of movements falling within the phee bounds, in a given gestational quadrant. The plotted errors are the standard errors of mean. The significance test was done using Fisher's exact test. (B) Each blue bar represents the percentage of movements falling within the twitter bounds, in a given gestational quadrant. The plotted errors are the standard errors of mean. The significance test was done using a two-tailed Fisher's exact test.

gestational quadrant. Error bars were calculated using the standard error of mean [two-tailed Fisher's exact test: general trend: p = .5e-04; quadrant 1 and 2: p = .31; quadrant 1 and 3: p = .4e-03; quadrant 1 and 4: p = .001; quadrant 2 and 3: p = .002; quadrant 2 and 4: p = .003; quadrant 3 and 4: p = .24].

For twitter syllables (in the infants the likelihood that a call within our 2-3 syllable criteria is a twitter call is .45), the increase in percentage of fetal movements matching our criteria persists (Figure 14B). In the figure, blue bars represent the percentage of movements falling within the twitter bounds in gestational quadrants. The error bars are standard errors of mean [two-tailed Fisher's exact test: general trend: p = .004; pairwise comparison: quadrant 1 and 2: p = .12; quadrant 1 and 3: p = .01; quadrant 1 and 4: p = .001; quadrant 2 and 3: p = .22; quadrant 2 and 4: p = .01; quadrant 3 and 4: p = .06]. It seems that late in gestation, the syllable features of infant calls (phee and twitter) are more evident in the orofacial movements of fetal marmosets.

An infant call is best defined by a combination of features and not duration or syllable number alone. Table 3 combines the duration and syllable number bounds we used for each call type. For each combination, the probability that a call within both ranges is of a particular type has also been listed. We see that taking duration and syllable number into account all calls except the trill call can be identified with greater accuracy (compare Table 1, 2 with Table 3). Here too, we only concentrated on phee and twitter calls (the two most discernable call types).

Binding these two features, we looked at fetal mouth movements that have both duration and syllable features of infant phee and twitter. For infant phee calls, the likelihood that a call within these ranges is a phee call is 0.97. Figure 15A shows that the percentage of calls fitting both duration and syllable requirements increases with increasing gestational age. The red bars indicate the percentage of fetal mouth movements captured by this sieve, per quadrant [two-tailed Fisher's exact test: general trend: p = .005; pairwise

	Phee Criteria 3.69 – 6.5s 5 – 9 syllables	**Twitter Criteria 0.96 – 1.5s 2 – 3 syllables**	**Trill Criteria 0.69 – 1s 1 syllable**	**Cry Criteria 1.17 – 3s 1 – 4 syllables**
Phee likelihood	**0.97**	0.02	0.13	0.16
Twitter likelihood	0	**0.86**	0.07	0.32
Trill likelihood	0	0.12	**0.51**	0.07
Cry likelihood	0.03	0	0.29	**0.46**

Table 3. Duration + syllable number definitions of infant call types. The column titles have the call type and duration + syllable number criteria of that particular call type. The values listed in the table are the likelihoods that a call meeting the duration + syllable number criteria is of a particular type. The highlighted values are the ones where the criteria and call type match (e.g. phee call criteria and the likelihood that a call within that range is a phee call).

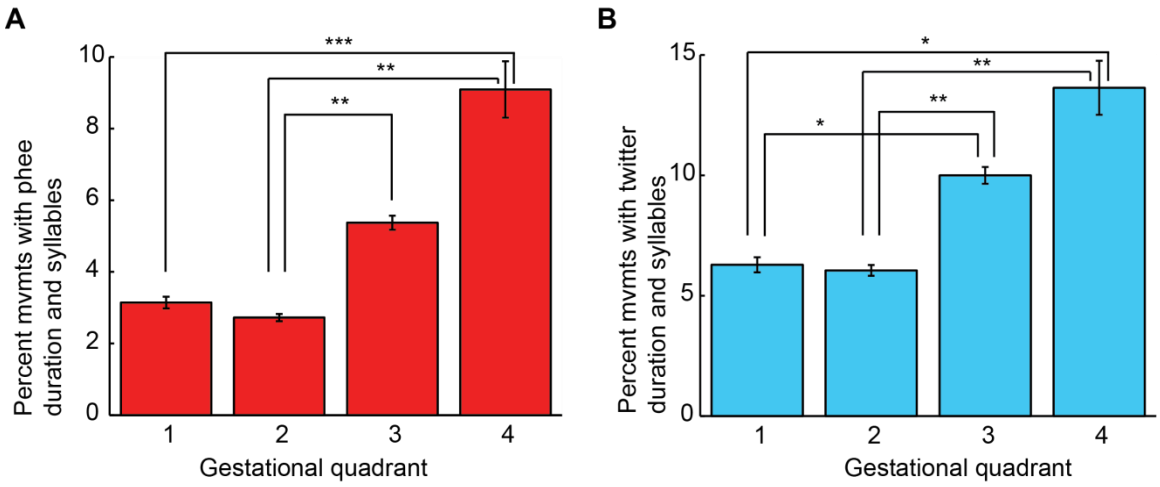


Figure 15. Percentage of fetal movements matching infant phee and twitter call duration and syllable number signatures. The tested period into four roughly equal quadrants — days 93-105 constitute the first gestational quadrant, days 106-118 the second quadrant, days 119-131 the third quadrant and days 132-148 the fourth quadrant period. (A) Each red bar represents the percentage of movements falling within the phee bounds, in a given gestational quadrant. The plotted errors are the standard errors of mean. The significance test was done using Fisher's exact test. (B) Each blue bar represents the percentage of movements falling within the twitter bounds, in a given gestational quadrant. The plotted errors are the standard errors of mean. The significance test was done using a two-tailed Fisher's exact test.

comparison: quadrant 1 and 2: p = .62; quadrant 1 and 3: p = .10; quadrant 1 and 4: p = .01; quadrant 2 and 3: p = .01; quadrant 2 and 4: p = .002; quadrant 3 and 4: p = .1].

For infant twitters too the percentage of fetal movements matching our criteria increases with advancing fetal age (in the infant the likelihood that a call within these ranges is a twitter call is 0.86). Figure 15B shows that the percentage of calls fitting both duration and syllable requirements increases with increasing gestational age. The blue bars indicate the percentage of fetal mouth movements that meet duration and syllable number requirements, per quadrant. [two-tailed Fisher's exact test: general trend: p = .005; pairwise comparison: quadrant 1 and 2: p = .84; quadrant 1 and 3: p = .04; quadrant 1 and 4: p = .02; quadrant 2 and 3: p = .007; quadrant 2 and 4: p = .006; quadrant 3 and 4: p = .21].

Therefore, we see that fetal movements increasingly match up with infant call types. As gestation progresses different features of the infant calls are bound together and a greater proportion of movements have multiple defining features of a call type. This is not restricted to a specific kind of call; for instance, the phee call which has a long duration and large syllable number. The trend was observed for two disparate call types; including the twitter call, which is of short duration and does not consist of as many syllables as the phee call. Our study is the first to show that aspects of non-human primate vocal behavior have a period of prenatal development.

Discussion

The labeling of primate infant behaviors as ‘innate’ partly exists due to difficulty in observing primates in the fetal stage of their lives. The improvement of ultrasound imaging techniques has opened up the possibility of observing fetal activity in considerable detail but the study of human fetuses poses yet another problem; it is hard to conduct detailed longitudinal observations of individual fetuses. We have performed ultrasound imaging on awake, pregnant marmoset monkeys (*Callithrix jacchus*), to

identify pre-cursors of infant vocal behaviors in the fetus. Not all parts of the fetal vocal apparatus can be observed non-invasively. However, same as with humans, monkey vocalizations are produced, in part, by tightly coordinated mouth movements (Ghazanfar and Rendall, 2008; Ghazanfar and Takahashi, 2014). By observing fetal marmosets from the first day the mouth is clearly visible to the day before birth we have studied how fetal movements relate to the postnatal mouth movements required for vocal production. We find that duration and syllable number features of two distinct calls produced by newborn marmosets are evident in the orofacial movements of fetuses. Figure 9 illustrates this finding. In addition, the percentage of vocalization-linked movements increases through gestation (Figures 13,14,15).

Our first step in this process was to characterize the vocalizations of newborn marmosets. In addition to cries, marmoset infants produce a number of adult-like vocal signals. These include phee calls, twitter calls and trill calls (Bezerra and Souto, 2008). The calls can be easily distinguished based on a combination of their spectral prints, temporal profiles, duration range and syllable number range (Figures 10). Unfortunately, we cannot use all of these features to inspect fetal movements. In order to identify vocalization-related movements in the fetus we first came up with criteria for defining the infant calls based on their duration and syllable numbers. The duration and syllable number ranges that best characterized these calls, and provided the best separation between them were chosen (Table 1, 2). We noticed that not all calls could be easily distinguished on the basis on these two features. This can be seen in Figure 11. Looking at duration, phee calls were the easiest to distinguish. They have the longest duration compared to the other 3 calls (Figure 11A, B). With syllable number too, the phee calls were the easiest to distinguish; they have a larger syllable number than the rest of the calls (Figure 11C, D). It looks like the fetal phee call precursors are the only call precursors that can be picked up with a satisfactory degree of certainty. For this reason we mainly concentrated on phee-call precursors. However, to show that our observations are not restricted to phee calls, we also looked for twitter call precursors. Among the four calls produced on day 1, the twitter call is the call that is most distinct from phee calls (Figure 10). A previous study from our lab shows that cries are immature forms of the phee call (Takahashi et al., in

press). The trill call is a short duration call with one syllable; it is similar in many ways to a single syllable of the phee call (compare Figure 10A and C). On the other hand, the twitter call is a short duration call with many syllables (though not as many syllables as the phee call). The twitter call cannot be clearly distinguished based on duration or syllable number alone (many cries fall into the mix) but they can be distinguished higher than chance when duration and syllable number are both considered (Table 3). For these reasons we picked the twitter as our contrast from the phee.

In infant marmosets (like in any other infants), the mouth musculature is used for more than just vocalizing. Feeding behaviors, for instance, also require well-developed mouth movements. These movements too must be developing in-utero. To track the development of fetal vocalization-related movements and exclude all the other movements, we filtered the fetal movements to include only those that fell within the bounds of vocalization duration range (upto 6.5s) and syllable number range (1-9 mouth movements in a single bout). These movements consisted of 93.33% of all fetal movements.

The developmental profiles of these vocalization-related movements show that they have increasing duration and syllable number with increasing fetal age (Figure 12). Infant marmoset calls are of varying durations but most are multi-syllabic, meaning each call is produced by a succession of open-and-shut mouth movements (Figures 10,11). Looking for specific features of infant phee and twitter calls, we saw that as gestation progresses there is an increase in the proportion of movements that have the characteristic duration range (Figure 13) and syllable number range (Figure 14) of the calls. The phee calls are of long duration (by our definition, 3.69-6.5s) and have many syllables (5-9). Given our previous finding that movements increase in duration and syllable number (Figure 12) with increasing fetal age, it is not surprising that a greater proportion of phee-like fetal movements are seen late in gestation (Figure 13A, 14A). From this finding alone we cannot conclude that precursors of orofacial mouth movements related to infant vocalizations develop in-utero. The twitter call, however, is a short duration call (by our definition, movements 0.69-1.5s in length); the proportion of movements within this

duration range also increases through gestation (Figure 13B). The proportion of movements matching twitter call syllable range (2-3) also rises with time (Figure 14B). The best definition of a call is one which includes both duration and syllable number. Looking at fetal movements that have the correct combination of duration and syllable number call features, there is an increase in the proportion of phee-like and twitter-like movements through gestation (Figure 15). This is the first evidence, that in a primate, infant vocal behavior has an extended period of development in the fetal stage.

Under normal conditions fetuses do not produce sounds. However, human fetuses in the last trimester have been heard to produce infant-like cries when air has entered the uterine cavity (Ryder, 1943). This suggests that motor development required for human vocal behavior also starts in-utero. There has been no study of the mouth movement trajectory in human fetuses but comparable developmental patterns have been observed for global motor development. For instance, motor behaviors are known to differentiate as gestation proceeds (de Vries et al., 1982; Roodenburg et al., 1991; DiPietro et al., 2001). By late fetal stage, the duration of movements (Groome et al., 1999), number of movements (Pillai and James, 1990) and cyclic motility patterns of movements (Robertson, 1987) are the same as those in neonates. Further, the results of our study and the development of human facial expressions show correspondences. Facial expressions resembling infant smiling, scowling, pouting, hiccupping and yawning are frequent at the end of gestation (Kurjak et al., 2003; Kurjak et al., 2005). A gradual progression from isolated facial movements toward fetal cry and laughter gestalts has been observed (Reissland et al., 2011). Therefore, human expressions and marmoset vocalizations do not appear de novo at birth; both have a protracted period of in-utero development. It is possible that same as with the marmosets, vocalizations human infants produce at birth have a long period of fetal development.

Our observations of fetal marmosets show three patterns in mouth movement maturation (1) The duration of fetal movements increase through gestation. (2) There is increasing organization of movements into multi-syllabic bouts. (3) Fetal mouth movements differentiate into different call types through gestation.

The increase in fetal mouth movement duration can be explained by a strengthening of mouth musculature with age but the increasing organization of movements into multi-syllabic bouts requires more than just increase in muscle strength. The development of mouth movement bouts can be compared to the development of limb movement bouts. In fetal rats, spontaneous limb movements have low levels of synchrony at the earliest observations periods and increased synchrony through development (Kleven et al., 2004). The result is that the probability of limbs moving in quick succession (defined by a criterion of 0.2 s or less) increases through gestation (Robinson et al., 2000). Similarly, we show that the occurrence of mouth movements in quick succession (defined by a criterion of 0.5 s or less) increases through gestation (Figure 12). These developments are necessary for the infant behaviors — vocalizations are produced by bouts of mouth movements and locomotion entails bouts of limb movements. The key difference is that spontaneous limb movements occur in bouts, which are nearly always composed of the activity of different limbs. Therefore, if you look at the activity of one limb alone, there is a refractory period after movement. Refractory periods after bursting activity are thought to be a general feature of spontaneous neural activity in developing circuits (Robinson et al., 2000). Mouth circuits could also have such a refractory period but the time interval could be much smaller, allowing for movement of the same muscle groups in short intervals. In rat fetuses, the spinal cord alone is sufficient to organize many aspects of spontaneous limb activity (Robinson et al., 2000; Kleven et al., 2004). This cannot be the case for mouth movements; instead, we propose the reticular formation region of the brainstem as a likely candidate. Experiments on cat fetuses show that this region has one of the earliest appearances of electrical activity (Gladkovich et al., 1980). It is also thought to control early fetal movements (Shuleikina and Gladkovich, 1965; Gladkovich et al., 1980). Human infants do not produce adult-like vocalizations at birth but they do cry; the brainstem is thought to be responsible for organizing crying behavior (Newman, 2007). In adult guinea pigs, the central pattern generator for rhythmical mouth movements is located in the reticular formation (Chandler and Tal, 1986; Nozaki et al., 1986; Hashimoto et al., 1989), and this area consists of premotor interneurons that are directly controlled by the mouth area of the motor cortex (Nozaki et al., 1986). In adult

primates, the brainstem reticular formation is known to be involved in vocal production; demonstrated by single-unit recordings and electrical stimulation (Jürgens and Hage, 2007).

Vocalizations of infant marmosets have characteristic durations and syllable numbers (Figure 11). Our results show that as gestation progresses fetal mouth movements develop these vocal signatures (Figures 13,14,15). One possibility is that duration and rhythmicity preferences are established through auditory feedback and motor mimicry of the auditory signal. Human fetuses can hear and prenatal learning of the mother's voice (DeCasper and Fifer, 1980), preference for native language over a foreign language (Moon et al., 1993) and preference to music their mother listened to during the last three months of pregnancy (Hepper, 1997) have been shown. However, we contend that prenatal development of infant vocal behavior is not dependent on auditory feedback. This is why: Winter, Handley, Ploog and Schott (Winter P and Schott, 1973) devocalized pregnant squirrel monkeys and isolated these females. On testing the vocal repertoire of the newborn infants they found that these infants produced all the call types produced by normal infants and structural aspects of the vocalizations were essentially normal; indicating that auditory input is not essential. Of course, it is possible that prenatal development in humans is dependent on auditory feedback but in non-human primates it is not (there is no experiment showing that non-human primates can hear in the womb). Nevertheless, we will still argue that prenatal development of vocal behavior does not depend on auditory feedback. In fact, it is more likely that the development of the auditory system piggybacks on prenatal vocalization-related motor patterns, which in-turn, is closely linked to somatosensory development.

The onset of sensory function in a young organism (birds and many mammal species) has an invariant sequence. The sequence being: somatosensory, vestibular, olfactory, auditory and visual (Gottlieb, 1971). During the critical period of mouth movement development the auditory system may not be online. Instead, motor development appears to be tied to somatosensory development. The following is one possible way in which the sensori-motor system develops: The first movements of a young fetus are generated by

spontaneous neural activity. These movements cause various types of interactions with the environment. Somatosensory information from these interactions begins to drive the organization of sensory networks. Motor networks receive this sensory information, and now motor areas begin to organize. Somatosensory and motor information are integrated based on the contingency of the information. As the network matures and the muscles get stronger, the complexity of movement increases [adapted from (Yamada et al., 2010); with the application of this mechanism their fetal simulations developed species-typical motor behaviors]. If we apply this principle to the development of duration and syllable number signatures of mouth movements, what we expect is that mouth area biomechanics (biomechanics of orofacial muscles, jaws, mandible joints), somatosensory feedback generated from orofacial contact and movement, and developing central pattern generators (CPGs) in the brain, all contribute towards the differentiation of mouth movements. Similar dynamics have been proposed for the development of infant speech (Kent, 1991). Different combinations of biomechanical values (muscle tension, jaw departure etc.), somatosensory stimulation and CPG network activation can generate movements of varying durations and syllable numbers. For instance, in premature human infants, for purposes of therapy, orofacial stimulation is used to increase the number of sucking mouth movements in a bout. The number of sucking movements entrains to the stimulation and this therapy leads to long-term improvements in sucking behavior (Barlow et al., 2008). Meaning that orofacial stimulation can alter movement patterns and drive the maturation of CPG networks. In cats repetitive electrical stimulation of the brain stem increases the speed of locomotion and switches the mode of locomotion (MacKay-Lyons, 2002). Stimulating locomotion CPGs at different frequencies also leads to movement bouts of different durations (Robinson and Kleven, 2005). In normal development preferred movement duration and syllable number ranges are likely those ranges that involve minimum energy expenditure. In humans, the syllable rate of natural languages seems to be determined largely by biomechanical solutions that minimize energy expenditure (Sorokin et al., 1980). Therefore, biomechanics of the body, somatosensory feedback and the developing neural networks work in tandem to set up motor patterns.

In our study we have concentrated on the development of mouth movements required for production of vocalizations. This is just one part of the vocal production apparatus. The generation of vocal signals requires coordination of respiration, mouth, larynx and tongue movements (Ghazanfar and Rendall, 2008; Ghazanfar and Takahashi, 2014). To understand how the entire behavior develops we can use the analogy of pecking behavior in chicks. Domestic chicks begin pecking at food grains etc. soon after hatching and this behavior has prenatal origins. The behavior consists of head-lunging, bill opening and closing, and swallowing. Full coordination is present only after birth and improves with postnatal practice, but the modules out of whose interaction pecking behavior emerges, develop in-utero. Head movement, observed as early as 3 days of embryonic age, is the first element to emerge. The movement is caused by amniotic contractions synchronized with fetal heartbeat. A day later, the bill begins to open and close when the un-hatched bird nods. Swallowing is the last to develop and is seen only 8-9 days after bill movement. It is thought that the repeated elicitation of head-lunging and bill movement by tactile stimulation from the yolk sac leads to increasing ties between the two, and the repeated elicitation of swallowing by amniotic fluid pressure due to bill opening adds the swallowing element to the pattern. Nevertheless, it is only when the chick is released from the confines of the egg and head-lunging is elicited by visual stimuli (e.g. the sight of food) that the full behavior is observed (Kuo, 1932a) (Kuo, 1932c; Kuo, 1932b; Lehrman, 1953).

The development of chick pecking behavior has many parallels to the development of vocal production. First, pecking behavior of chicks and marmoset vocal production, develop in-utero. Second, pecking behavior does not develop as a pecking *pattern.* The different modules of the pecking behavior develop at different rates and via different mechanisms. Likewise, the different movements required for vocal production have different developmental trajectories. For example, mouth movement and respiratory coupling is necessary for vocal production (Sakamoto et al., 1997; Jürgens, 2002; Smotherman et al., 2006). In the first half of fetal life, breathing movements are independent of other movements and mouth movements are not seen when the fetus is breathing (Piontelli, 2010). It is only late in fetal life that the respiration CPG couples

with other movement CPGs (Piontelli, 2010). [The reticular activating system is known to mediate motor coordination of respiration (Zeskind and Lester, 2001; LaGasse et al., 2005), in addition to larynx and the mouth]. Third, the influence of fetal environment is crucial for the development of behavior. Tactile stimulation from the yolk sac and pressure from amniotic fluid, aid the development of pecking behavior. In Chapter 1 we hypothesize that decrease in uterine space with increasing fetal size is part of what drives the differentiation of mouth and head motor modules. Infant marmosets can then use these parts for different functions (mouth for feeding and vocalizing; head for orienting). Fourth, the behavior of the fetus itself drives the self-organization of behavior. In pecking behavior, bill opening causes pressure from the amniotic fluid and elicits swallowing — ultimately tying the two elements together. Human fetuses are known to touch their mouth area with their hands, increasingly so through gestation (de Vries et al., 1982; DiPietro et al., 1996; Kurjak et al., 2005). We did not observe the monkey fetuses do the same but we observed that they repeatedly put their tails in their mouth. The mouth area is rich in tactile receptors ((Mori and Kuniyoshi, 2010). These kinds of stimulations along with sensations caused by movement probably excite the mouth muscles and serve to bind these parts together through hebbian neural mechanisms. Lastly, 'pecking' behavior in the womb does not look like postnatal pecking behavior. The chick has to be released from the confines of the egg, and visual stimuli are needed for the full pecking behavior to be seen. Similarly, for the vocal apparatus to act together and produce sounds, air is required (Ryder, 1943). In the case of vocal production, auditory stimuli can elicit the behavior. Therefore, fetal behaviors are the foundations for postnatal behaviors and, (like postnatal behaviors) they emerge in a non-linear fashion from the organization of the current developmental phase of an organism and interactions between this developing organism and its current environment.

The prediction would then be that perturbations to body morphology, brain structure or changes in fetal environment lead to a different developmental outcome. We see this in primate orofacial development — changes in temperature due to the mother running a fever can cause defects in the fetus; the severity and incidence of defects depend largely on the developmental stage of the fetus at the time of exposure and on the amount of

temperature change. This has been observed in monkeys and humans (Graham and Edwards, 1998). Human fetal defects include abnormal muscle tone, cleft lip and/or palate, undersized jaws and conditions where the center of the face develops more slowly than the eyes, forehead, and lower jaw (Chambers and Jones, 1997). Orofacial clefting can also be caused by maternal cigarette smoking, drinking and use of steroids (Eppley et al., 2005). These defects caused by environmental perturbations prove that fetal orofacial development does not follow an invariable or inevitable course; the sequence and outcome of development is probable rather than certain.

Mouth movement behavior and therefore infant vocal behavior is not 'innate'. At every point in development, constraints set-up by the current brain organization, current morphology (e.g. physical limits of muscle fibers and/or joints) and current environment (e.g. amniotic fluid and uterine space) drive the self-organization of behavior. Ignoring the fact that fetal behaviors lay the foundation for later, comparatively mature infant behaviors has led to the erroneous conclusion that behaviors present at birth are 'innate'.

Methods

Subjects

We tracked four pregnancies and a total of nine fetuses. This included one pregnancy of a single baby, two pregnancies of twins, and one pregnancy of quadruplets.

Of the nine fetuses, seven survived (two of the quadruplets were stillborn). There were no obvious differences in the sizes of the singleton, twins, or quadruplets at birth (including the stillborn infants). Our veterinary doctor performed necropsies on the stillborn animals and concluded that both were well developed and looked to be at term.

The four studied pregnancies were of two adult marmosets - three pregnancies of one female and one pregnancy of the other. Both animals were part of marmoset social groups housed at the Developmental Neuromechanics and Communication Laboratory, Princeton University. The two pregnant females came from different social groups and

lineages. Animals were fed once daily with standard commercial chow (Harlan Marmoset Diet #TD.08482.PWD) supplemented with fresh fruits and vegetables. Additional treats (typically peanuts, cereal, dried fruit, raisins and marshmallows) were given during experimental sessions. The animals had ad libitum access to water. The colony room was maintained at a temperature of approximately 27°C and 50-60% relative humidity, with a 12L:12D light cycle. Previous to conducting the experiment, all animals were trained to leave their home cage in a transport box and familiarized with the testing room and testing equipment. All experimental sessions were in daylight hours - between 1400 and 1800. All procedures were in compliance with guidelines of the Princeton University Institutional Animal Care and Use Committee.

Ultrasonography technique

Ultrasonography tests were performed using a GE Voluson i Ultrasound machine. The testing method used was a modified version of the procedure developed by S.D. Tardif and colleagues (Jaquish et al., 1995). Each examination was conducted by two experimenters — one gently restrained the animal while the other carried out the ultrasound procedure. The animals were not anesthetized for the examination. Instead, they were trained to accept gentle restraint and were rewarded with treats at the end of the testing period. We routinely tested all our adult females for pregnancy. Early pregnancy was detected by assessing the morphology of the uterus using the method described by Jaquish and colleagues (Jaquish et al., 1995). Once pregnancy was confirmed, ultrasounds were conducted every two weeks to monitor the development of the fetus. When the fetal skull became clearly visible, gestational age could be estimated by measuring the biparietal diameter of the skull (Jaquish et al., 1995). At this point, we started monitoring the fetus 2-3 times a week. As soon as the fetal face was clearly visible (usually around 95 days gestational age), examinations were conducted every weekday. A total of 64 sessions were recorded (14 sessions from the first pregnancy; 17 sessions from the second; 17 sessions from the third; 16 sessions from the fourth pregnancy). The examination was terminated if the animal showed significant resistance. A typical experiment ranged from 15-40 minutes. During these examinations the probe was manipulated to ensure that a frontal view of the face was attained at all times. In

cases where there was more than one fetus, two fetuses were selected at random and each fetus was observed for 15-20 minutes. The ultrasound videos were captured at a frame rate of 29.97 frames per second and written on to DVD for later analysis. All exams were performed between 14:00 and 18:00 hours and the procedure was repeated until birth (usually around 146 days gestational age).

Recording infant vocalizations

Audio and video recordings of infant vocalizations were made ~24 hours after birth. The infant was separated from the parents and taken to a 2.5 m X 2.5 m room with walls covered in sound attenuating foam. For the audio recordings, the infant was placed on a layer of foam in a metal transfer cage. Once the subject was in place, the experimenter turned on a digital recorder (ZOOM H4n Handy Recorder) positioned directly in front of the testing cage at a distance of 0.76 m and left the room for a period of 5 minutes. For the video recordings one experimenter held the baby and a second experimenter acquired videos of the infant face using a hand-held SONY video recorder. This was done to observe facial movements during vocalizations. The frame rate of the infant videos matched that of the ultrasound videos — both were 29.97 frames per second.

Exemplars — infant call types

The image panels were generated by first using a custom-made MATLAB® program to split chosen video clips into the component frames and then selecting those frames which best exemplified the movement.

Extracting the audio signal from the chosen clips and reducing background noise using the Adobe Audition software produced the spectrograms.

Temporal profiles of the accompanying mouth movements were generated using a custom made MATLAB® program which allowed us to go through the movement clip frame-by-frame, mark the lower and upper jaw line and calculate the pixel distance between the two. When there was no mouth movement, this distance is 0, while a nonzero value indicates that the mouth was open. The profiles were then z-scored to

remove amplitude information. The amplitude information was removed because we cannot reliably gather amplitude measurements using ultrasonography. The generated profiles were minimally smoothed using the cubic smoothing spline (csaps) function in MATLAB. The smoothing parameter was .999 (MATLAB R2012a)

Analysis of infant vocalizations — audio recordings

The audio recordings were processed using Adobe Audition software. Spectrograms of the audio signals were used to categorize the infant vocalizations. The infants produced calls of 4 major types (~92% of P1 vocalizations). Of these, 20% were identified as phee calls, 18% twitter calls, 13trill calls and 42% cries. The number of syllables in each call ranged from 1 (mostly trills) to 9 syllables (phees and cries). Onset-offset gaps of 500ms or longer indicated separate calls, whereas gaps shorter than 500 ms indicated syllables from the same call. The median duration and number of syllables of each call type were calculated.

Analysis of infant vocalizations — video recordings

Video analysis was done using the Adobe Premiere Pro software, which enabled us to screen the videos frame-by-frame. The first video frame where the jaws separated counted as the onset of mouth movement and the first video frame where the jaws came back together was the offset of the movement. Individual mouth movements were considered to be part of the same movement unit if they were separated by 500ms or less (~15 frames). Here too median call duration and syllable number of each call type were calculated.

Equating the audio and video duration measures

A mismatch was seen in the audio and video duration measures for the same call. The video measures were consistently longer than audio measures. This could be a due to a technical mismatch, due to the dynamics of call production (it is possible that the mouth starts opening a little before the sound is actually heard), or a combination of the two. To compensate for the duration difference, we calculated the difference between the video and audio duration and grouped the differences by call type. We calculated the median

duration discrepancy for the phee and the 90 percentile discrepancies for the other call types. The median value was added to the audio duration of all phee calls and the 90 percentile values were added for the other call types. The number of syllables remained the same across the audio and video measures; no adjustment was required.

Kernel smoothing density estimates for infant call durations

The probability density estimates for the duration distributions of infant calls were calculated using the MATLAB ksdensity function. For the violin plot (Figure 11A) we used the 'Violin plot based on kernel density estimation' routine created by Holger Hoffmann (MATLAB file exchange #45134).

Syllable number distribution for infant calls

The probability density estimates for the syllable number distributions of infant calls were calculated using the MATLAB ksdensity function. The boxplots for syllable number ranges distributions (Figure 11C) we used the MATLAB boxplot function.

Identifying fetal mouth movements

Ultrasound video analysis was done using the Adobe Premiere Pro software. Since the animals were not anesthetized for the ultrasound procedure, both the mother and the fetus could move, causing us to lose sight of the fetal face from time to time. Therefore, we first identified segments of the video during which the fetal face was clearly visible. Within these segments, we scored for mouth movements. A movement counted as a mouth movement when there was a clear separation of the upper and lower jaws (Figure 10B image panels, mouth openings marked in red and blue). The first video frame where the jaws separated counted as the beginning of mouth movement and the first video frame where the jaws came back together was the end point of the movement. If we lost focus of the face before the jaws came back together, the last frame of observation counted as the end of the movement. Individual mouth movements were considered to be part of the same movement unit if they were separated by 500ms or less (~15 frames). The 500ms criterion was justified by the bimodal structure of the inter-syllable interval distribution of the vocal output of marmoset neonates. In their calls the 500ms threshold separates the

first mode of the distribution (representing the interval between syllables within a single call) from the second mode that represents the interval between the offset of the last and onset of the first syllables between two calls (DiMattina and Wang, 2006). Henceforth, when we refer to mouth movements we are speaking of movement units.

Exemplars — fetal mouth movements matching phee and twitter calls

First we picked two fetal exemplars (Figure 10B), one that matched the phee duration and syllable number criteria and one that matched the twitter duration and syllable number criteria. The fetal image panels was generated by first using a custom-made MATLAB® program to split the chosen video clips into its component frames and then selecting those frames which best exemplified the movement. The red and blue lines on the images indicate mouth opening. The temporal profile was generated using a custom made MATLAB® program which allowed us to go through the movement clip frame-by-frame, mark the lower and upper jaw line and calculate the pixel distance between the two. When there was no mouth movement, this distance was 0, while a nonzero value indicated that the mouth was open. The amplitude of the movement did not have any bearing on the analysis. The generated profiles were smoothed using the cubic smoothing spline (csaps) function in MATLAB. The smoothing parameter was .999 (MATLAB version R2012a).

Duration and syllable number of fetal movement

Mouth movement durations were calculated by subtracting a movement's offset time from its onset time. Syllable numbers were calculated by counting the number of mouth movements in each movement unit. We were only interested in vocalization-related movements so we filtered out all other types of fetal movement. This was done by only including fetal movements within the range of vocalization durations (upto 6.5 seconds) and vocalization syllable numbers (1-9 syllables).

Contour plots for the fetal duration and syllable number (Figure 4)

Each black dot on the graph represents the median movement duration of one session; the black squares and black triangles show the upper 90 and lower 10 percentile points,

respectively, for each session. The curves were generated by first calculating the optimal degree for polynomial fitting according to Akaike's information criterion (AIC) (Akaike, 1981b), for the median values. The same polynomial fitting degree was used for the other two curves. The contour plots for median durations and syllable numbers were made using MATLAB routine 'Kernel density estimation' by Zdravko Botev.

Gestational Quadrants

We divided the tested period into four roughly equal quadrants — days 93-105 constitute the first gestational period, days 106-118 the second period, days 119-131 the third period and days 132-148 the fourth gestational period.

Percentage of movements matching phee and twitter durations

Mouth movement durations were calculated and results from all pregnancies were pooled. The percentage of movements matching the phee and twitter duration criteria were calculated by counting the number of movements that fell within the phee duration range (3.69-6.5s) and twitter duration range (0.69-1.5s) in each gestational quadrant, then dividing these numbers by the total number of calls in that quadrant. The error bars are standard errors of mean. To test for the significance of the trend we used a two-tailed Fisher's exact test. The calculations were done using the MATLAB routine 'MyFisher' by Giuseppe Cardillo (MATLAB file exchange #26883).

Percentage of movements matching phee and twitter syllable numbers

The numbers of mouth movements in each movement unit were counted and compiled for all pregnancies. The percentage of movements matching the phee and twitter syllable criteria were calculated by counting the number of movements that fell within the phee syllable range (5-9) and twitter duration range (2-3) in each gestational quadrant, then dividing these numbers by the total number of calls in that quadrant. The error bars are standard errors of mean. To test for the significance of the trend we used a two-tailed Fisher's exact test. The calculations were done using the MATLAB routine 'MyFisher' by Giuseppe Cardillo (MATLAB file exchange #26883).

Percentage of movements matching phee and twitter durations plus syllable numbers

These percentages were calculated in the same way as above, except we counted the number of movements that met duration and syllable number criteria. Statistical significance was tested using Fisher's exact test, in the same way as before.

General Discussion

The turning point in the history of man was when he acquired the use of verbal concepts and could organize his experiences in a common pool. This evolutionary advantage has allowed us to override conventional natural selection and shape our own future, an act that has not been possible for any animal before us. As unique as is it may be, it is important to remember that our vocal behavior is a product of evolution and has it's biological roots in animal communication. Nonhuman primates in particular, have many likenesses to humans — in brain structure and physical attributes (Burman and Rosa, 2009; de la Mothe et al., 2012) and, social behaviors (Takahashi et al., 2013). Given these similarities, precursors of human speech must be present in the vocal communication of other primates. One behavioral difference, leading us to think of primate communication as entirely separate from human communication, is that newborn monkeys produce calls that are structurally comparable to adult calls (Bezerra and Souto, 2008). Until now there has been no careful study of their prenatal development. Nonhuman primates are born precocial and develop at a much faster rate than humans. The presence of adult-like calls at birth could be due to a larger part of vocal development occurring in-utero, whereas; in us, prenatal development takes us only to the point of producing cries. Our prolonged developmental time has probably allowed us to add additional layers to our communication system. This does not mean that the basic principles of development differ between them and us. This thesis has aimed to uncover basic principles of prenatal vocal development.

To study how fetal movements relate to the postnatal mouth movements required for vocal production, we have performed ultrasound imaging on awake, pregnant marmoset monkeys (*Callithrix jacchus*), from the day the fetal mouth is first visible to the last day of birth. Not all parts of the vocal apparatus can be visualized via ultrasonography. As with humans, however, monkey vocalizations are produced in part by tightly coordinated mouth movements (Ghazanfar and Takahashi, 2014); these movements can be observed in the fetus. Our longitudinal study has allowed observation of the entire developmental trajectory of mouth movements. The non-invasive nature of this study has limitations (for

instance, it is challenging to tease apart the underlying mechanisms without manipulation) but, it has enabled observation of the fetus in it's natural environment.

After their first appearance, mouth movements gradually increase in quantity and later decrease after hitting their peak further in gestation. By the end of gestation, mouth movements (and other observed movements) almost completely die out (Figure 2). This inverse-U shaped profile has been observed across different species (rats, ducks, pigeons, chicks and even humans), for many types of movements (general movement, limb movement, jaw movement, eye saccades) (Barron and Donald, 1941; Oppenheim, 1974; Gottlieb, 1976; Smotherman and Robinson, 1988; Roodenburg et al., 1991). The decline in movement quantity before birth seems to be evolutionarily conserved. The most common explanation for this decline is increasing physical constraint from the uterine environment. We have evidence suggesting that decreasing uterine space does not account for all of the change (Figure 3); inhibition from the developing brain possibly adds to the decline in movement quantity.

Fetal mouth movements first occur with head movements. The mouth and head gradually stop moving as a unit and by the end of gestation, they are moving independently (Figure 5A, Figure 7). Other research groups (Mori and Kuniyoshi, 2010; Mori, 2012; Yamada and Kuniyoshi, 2012) have modeled the motor development of human and zebrafish embryos. The, fetal models were initialized with different physical attributes (size, musculoskeletal configuration and tactile receptor distribution), but the models enlisted identical neural circuits. Spontaneous movements dictated by the physical properties of the model led to differential neural organization, which in turn influenced movement. In both mock-ups, one motor module initially controlled multiple body parts (e.g. a single unit controlling the head and trunk). As development progressed these large units subdivided into more precise ones (e.g. separate control of head and trunk); leading to organized, species-typical movement patterns. A similar trend has been seen in experimental work on chicks, where the wings and legs are coupled early in gestation and decouple before birth (Provine, 1980; Bradley, 1999). This phenomenon has not been described in humans, but that is largely due to the considerable ethical and practical

boundaries to setting up a longitudinal prenatal study. We can, however infer that such a developmental trajectory exists. Evidences from aborted fetuses suggest that mouth and head parts are controlled by the same motor unit in early points of fetal life. (Hooker, 1936; Fitzgerald and Windle, 1942). In aborted fetuses, Hooker and colleagues found that a light touch to the perioral region elicited contractions of the neck muscles; Fitzgerald and Windle found that stimulation of the maxillary region of the mouth caused contralateral head flexions (Fitzgerald and Windle, 1942). It is therefore likely that the developmental course we describe is maintained in humans. We posit that inhibition from the developing brain, differing uterine constraint on the mouth and head, and differential tactile receptor distribution on the fetal body, all serve to organize the connections between mouth and head motor modules, via anti-hebbian rules of neural pruning (Figure 8).

In regard to prenatal mouth movements in the marmoset, we observed that signature features of marmoset infant calls emerge prenatally as distinct patterns of orofacial movements: Late in gestation, the duration and syllable number of phee and twitter calls — calls produced by marmoset monkeys on postnatal day 1 — are present in the mouth movements of fetal marmosets. This is evidence that aspects of vocal behaviors in marmosets have a period of prenatal development. Under normal conditions fetuses do not produce sounds but human fetuses in the last trimester have been heard to produce infant-like cries when air has entered the uterine cavity (Ryder, 1943). This suggests that motor abilities required for human infant vocalizations also develop in the womb. There has been no study centered on vocal development in the prenatal stage of humans but certain features of general motor activity are maintained between late fetal and neonatal periods; including, duration of movements (Groome et al., 1999), number of movements (Pillai and James, 1990) and, cyclic motility patterns (Robertson, 1987). Progression toward complexity has been observed in other motor behaviors, such as the developmental change from isolated facial movements toward infant-like cry and laughter facial gestalts observed in human fetuses (Reissland et al., 2011). In fact, by the end of gestation, facial expressions resembling infant smiling, scowling, pouting, hiccupping and yawning are frequent (Prechtl, 1985; Kurjak et al., 2005). In both

vocalization and expression-related fetal movements there is increased coordination between certain muscle groups with age. In marmosets, there is increased coordination of mouthparts leading to the emergence of infant call signatures. In humans, increased coordination of facial movements leading to the emergence of infant-like facial expressions.

We have therefore shown: (1) Aspects of vocal behaviors in marmosets have a period of prenatal development. (2) Mouth and head modules become increasingly differentiated through gestation; allowing the infant to use these regions for separate functions (mouth for feeding and vocalizing; head for orienting). (3) Mouth musculature becomes increasingly coordinated through gestation; allowing the infant to produce vocalizations on the first day of birth. Further, for each of these findings, analogous studies point to homologous development in humans.

It has been difficult to devise uniform theoretical schemes of the mechanics underlying development as it does not take place in a linear uniform fashion and, for lack of longitudinal studies in individual subjects. For a flavor of the non-linearity of development: Movement patterns involving multiple body parts differentiate into isolated movements (Coghill, 1929; Bradley, 1999); on the other hand, isolated movements are present at the earliest stages of fetal development (Kuo, 1939; Windle, 1944; Narayanan et al., 1971); to add to this, irregular, uncoordinated, jerky movements involving multiple body parts can be seen late in gestation, even after coordinated movements have developed (Hamburger and Oppenheim, 1967; Bekoff, 2001). These observations have led to debates about whether movements begin as part of a ‘system of total integration,’ which over time develops into independent local patterns — the ‘global-to-local’ argument (Coghill, 1929) or whether movements begin as local responses that later integrate into coordinated action patterns — the ‘local-to-global’ argument (Windle, 1944; Narayanan et al., 1971). Combining all our results, we see increasing differentiation and increasing coordination of motor modules.

If we start with what an infant can do at birth we do not need to pick between the ‘global-

to-local' and 'local-to-global' camps; or, as is customary these days, we do not need to avoid these distinctions entirely. Infants use their mouth and head to perform separate functions (mouth for feeding and vocalizing; head for orienting). The prenatal relationship between these two parts is different from the decoupled character of their postnatal versions. What we have found is, mouth and head movements occur together early in gestation but gradually decouple and move independently by the end of gestation (Figure 5A) — i.e. increased differentiation, a progression from 'global to local'. We have also found that mouth movements increase in duration and each bout of movement consists of more constituent movements later in gestation (Figure 12). Additionally, greater incidence of movements with signature traits of infant vocal signals is seen late in gestation (Figures 13,14,15) — i.e. increased coordination of mouth muscles, a progression from 'local to global'. What we could consistently observe with ultrasound imaging were only the jaws of the fetus. Since the infant can vocalize right at birth, it is very likely that other parts of the vocal apparatus (e.g. tongue and lips) also integrate with the jaw movements. In summary, for the infant to be able to vocalize at birth, 'global-to-local' and 'local-to-global' motor development is necessary.

The fact that nonhuman primates produce — in addition to cries — adult-like calls on the day of birth, has led to a commonly held view that their vocalizations are 'innate'. Our studies show that aspects of vocal behaviors in marmosets have a period of prenatal development. We have evidence suggesting that it is the dynamic interactions between the developing body and brain, and between the body and its environment, that drive the self-organization of vocalizations in the fetal stage. This means that primate vocal development does not follow an invariable or inevitable course; the sequence and outcome of vocal development is probable rather than certain. Perturbations to body morphology, brain structure or changes in fetal environment can lead to a different outcome. This refutes the 'innate' argument and even invites a role for prenatal influences on clinical syndromes.

A study on infant vocal development in marmosets (Takahashi et al., in press) shows that infant cry-like vocalizations transform into mature calls at around 2 months of age.

Measuring their vocal interactions with parents, we see that contingent vocal responses from parents influence the timing of transition from cries to mature calls. Thus, in both humans and marmoset monkeys, early vocalizations are the foundation for later, more mature vocalizations and, vocal transformation is influenced by social feedback.

At cursory glance, human speech may seem entirely different from that of other primates. On closer examination, the building blocks of our system can be found in other primates. The process of evolution is not unlike the process of development. The abilities of our closest extant relatives are the foundation for the later, more advanced abilities of man.

References

Akaike H (1981a) Citation Classic - a New Look at the Statistical-Model Identification. Cc/Eng Tech Appl Sci:22-22.

Akaike H (1981b) Citation classic - A new look at the statistical-model identification. Current/Contents/Engineering Technology & Applied Sciences:22-22.

Arbib Michael A, Liebal K, Pika S (2008) Primate Vocalization, Gesture, and the Evolution of Human Language. Current Anthropology 49:1053-1076.

Barlow S, Finan D, Lee J, Chu S (2008) Synthetic orocutaneous stimulation entrains preterm infants with feeding difficulties to suck. Journal of Perinatology 28:541-548.

Barron BY, Donald H (1941) The functional development of some mammalian neuromuscular mechanisms. Biological reviews of the cambridge philosophical society 16:1-33.

Bekoff A (2001) Spontaneous embryonic motility: an enduring legacy. International Journal of Developmental Neuroscience 19:155-160.

Bekoff A, Lau B (1980) Interlimb coordination in 20-day-old rat fetuses. Journal of Experimental Zoology 214:173-175.

Berivoj Miskovic OVMP (2011) Fetal Neurobehavior in Normal and High-Risk Pregnancy. Donald School Journal of Ultrasound in Obstetrics and Gynecology 5:101-106.

Bezerra BM, Souto A (2008) Structure and usage of the vocal repertoire of Callithrix jacchus. International Journal of Primatology 29:671-701.

Blankenship AG, Feller MB (2010) Mechanisms underlying spontaneous patterned activity in developing neural circuits. Nature reviews Neuroscience 11:18-29.

Blumberg MS, Marques HG, Iida F (2013) Twitching in Sensorimotor Development from Sleeping Rats to Robots. Current Biology 23:R532-R537.

Bradley NS (1999) Transformations in embryonic motility in chick: kinematic correlates of type I and II motility at E9 and E12. Journal of neurophysiology 81:1486-1494.

Burman KJ, Rosa MGP (2009) Architectural subdivisions of medial and orbital frontal cortices in the marmoset monkey (Callithrix jacchus). The Journal of comparative neurology 514:11-29.

Chambers CD, Jones KL (1997) Hyperthermia in pregnancy: A prospective cohort study. Teratology.

Chandler SH, Tal M (1986) The effects of brain stem transections on the neuronal networks responsible for rhythmical jaw muscle activity in the guinea pig. The Journal of neuroscience : the official journal of the Society for Neuroscience 6:1831-1842.

Cheung KL, Tunik E, Adamovich SV, Boyd (2014) Neuroplasticity and virtual reality. In: Virtual Reality for Physical and Motor Rehabilitation (Weiss PL, Keshner EA, Levin MF, eds), pp 5-24. New York, NY: Springer New York.

Chiel HJ, Beer RD (1997) The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment. Trends in Neurosciences 20:553-557.

Coghill GE (1929) Anatomy and the problem of behavior: MacMillan Co., New York.

de la Mothe LA, Blumell S, Kajikawa Y, Hackett TA (2012) Cortical connections of auditory cortex in marmoset monkeys: lateral belt and parabelt regions. Anatomical record (Hoboken, NJ : 2007) 295:800-821.

de Vries JIP, Visser GHA, Prechtl HFR (1982) The emergence of fetal behaviour. I. Qualitative aspects. Early Human Development 7:301-322.

de Vries JIP, Visser GHA, Prechtl HFR (1985) The emergence of fetal behaviour. II. Quantitative aspects. Early Human Development 12:99-120.

DeCasper A, Fifer W (1980) Of human bonding: newborns prefer their mothers' voices. Science 208:1174-1176.

Decker JD, Hamburger V (1967) The influence of different brain regions on periodic motility of the chick embryo. The Journal of experimental zoology 165:371-383.

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

DiPietro JA (2008) Prenatal Development. Encyclopedia of Infant and Early Childhood Development 2:604-614.

DiPietro JA, Hodgson DM, Costigan KA, Hilton SC, Johnson TRB (1996) Fetal Neurobehavioral Development. Child Development 67:2553-2553.

DiPietro JA, Bornstein MH, Costigan KA, Pressman EK, Hahn C-S, Painter K, Smith BA, Yi LJ (2001) What does fetal movement predict about behavior during the first two years of life? Developmental psychobiology 40:358-371.

Encha-Razavi F, Sonigo P (2003) Features of the developing brain. Childs Nervous System 19:426-428.

Eppley BL, van Aalst JA, Robey A, Havlik RJ, Sadove AM (2005) The Spectrum of Orofacial Clefting. Plastic and Reconstructive Surgery 115:101e-114e.

Fitzgerald JE, Windle WF (1942) Some observations on early human fetal movements. The Journal of Comparative Neurology 76:159-167.

Földiák P (1990) Forming sparse representations by local anti-Hebbian learning. Biological cybernetics 64:165-170.

Ghazanfar AA, Rendall D (2008) Evolution of human vocal production. Current Biology 18:R457-R460.

Ghazanfar AA, Takahashi DY (2014) Facial expressions and the evolution of the speech rhythm. Journal of cognitive neuroscience 26:1196-1207.

Gladkovich NG, Leontovich TA, Shuleĭkina KV (1980) [Quantitative morphologic characteristics of developing brain stem reticular formation neurons]. Neiřofiziologiia = Neurophysiology 12:53-61.

Gottlieb G (1971) Development of species identification in birds: An inquiry into the prenatal determinants of perception: University of Chicago Press.

Gottlieb G (1976) Conceptions of prenatal development: Behavioral embryology. Psychological Review 83:215-234.

Graham JM, Edwards MJ (1998) Teratogen update: gestational effects of maternal hyperthermia due to febrile illnesses and resultant patterns of defects in humans. Teratology 58:209-221.

Granmo M, Petersson P, Schouenborg J (2008) Action-based body maps in the spinal cord emerge from a transitory floating organization. The Journal of neuroscience : the official journal of the Society for Neuroscience 28:5494-5503.

Groome LJ, Swiber MJ, Holland SB, Bentz LS, Atterbury JL, Trimm RF (1999) Spontaneous motor activity in the perinatal infant before and after birth: Stability in individual differences. Developmental psychobiology 35:15-24.

Grossman RC, Hattis BF, Ringel RL (1965) Oral tactile experience. Archives of Oral Biology 10:691-IN636.

Gruest N, Richer P, Hars B (2004) Emergence of long-term memory for conditioned aversion in the rat fetus. Developmental psychobiology 44:189-198.

Hadders-Algra M (2007) Putative neural substrate of normal and abnormal general movements. Neuroscience and biobehavioral reviews 31:1181-1190.

Hamburger V (1963) Some Aspects of the Embryology of Behavior. The Quarterly Review of Biology 38:342-342.

Hamburger V, Oppenheim R (1967) Prehatching motility and hatching behavior in the chick. The Journal of experimental zoology 166:171-203.

Hammerschmidt K, Fischer J (2008) Constraints in primate vocal production. In, pp 93-119. Cambridge, MA: MIT Press.

Hashimoto N, Katayama T, Ishiwata Y, Nakamura Y (1989) Induction of rhythmic jaw movements by stimulation of the mesencephalic reticular formation in the guinea pig. The Journal of neuroscience : the official journal of the Society for Neuroscience 9:2887-2901.

Hayat T, Nihat A, Martinez-Biarge M, McGuinness A, Allsop J, Hajnal J, Rutherford M (2011) Optimization and initial experience of a multisection balanced steady-state free precession cine sequence for the assessment of fetal behavior in utero. American Journal of Neuroradiology 32:331-338.

Hepper PG (1997) Memory in utero? Developmental Medicine and Child Neurology 39:343-346.

Hikishima K, Sawada K, Murayama AY, Komaki Y, Kawai K, Sato N, Inoue T, Itoh T, Momoshima S, Iriki A, Okano HJ, Sasaki E, Okano H (2013) Atlas of the developing brain of the marmoset monkey constructed using magnetic resonance histology. Neuroscience 230:102-113.

Hooker D (1936) Early fetal activity in mammals. Yale Jour Biol and Med 8:579-602.

Izard CE (1994) Innate and universal facial expressions: evidence from developmental and cross-cultural research. Psychological bulletin 115:288-299.

Jaquish CE, Toal RL, Tardif SD, Carson RL (1995) Use of ultrasound to monitor prenatal growth and development in the common marmoset (Callithrix jacchus). American Journal of Primatology 36:259-275.

Jürgens U (2002) Neural pathways underlying vocal control. Neuroscience & Biobehavioral Reviews 26:235-258.

Jürgens U, Hage SR (2007) On the role of the reticular formation in vocal pattern generation. In, pp 308-314.

Kent RD, Mitchell P.R., Sacier M. (1991) Evidence and role of rhythmic organization in early vocal development in human infants. In: The development of timing and temporal organization in coordinated action (Fagard J WPH, ed), pp 135-150.

Kleven GA, Lane MS, Robinson SR (2004) Development of interlimb movement synchrony in the rat fetus. Behavioral neuroscience 118:835-844.

Kuo ZY (1932a) Ontogeny of embryonic behavior in Aves. I. The chronology and general nature of the behavior of the chick embryo. Journal of Experimental Zoology 61:395-430.

Kuo ZY (1932b) Ontogeny of embryonic behavior in aves. IV. The influence of embryonic movements upon the behavior after hatching. Journal of Comparative Psychology 14:109-122.

Kuo ZY (1932c) Ontogeny of embryonic behavior in aves. II. The mechanical factors in the various stages leading to hatching. Journal of Experimental Zoology 62:453-487.

Kuo ZY (1939) Total pattern or local reflexes. Psychological Review 46:93-122.

Kurjak A, Stanojevic M, Andonotopo W, Scazzocchio-Duenas E, Azumendi G, Carrera JM (2005) Fetal behavior assessed in all three trimesters of normal pregnancy by four-dimensional ultrasonography. Croatian Medical Journal 46:772-780.

Kurjak A, Azumendi G, Veček N, Kupešic S, Solak M, Varga D, Chervenak F (2003) Fetal hand movements and facial expression in normal pregnancy studied by four-dimensional sonography. Journal of Perinatal Medicine 31:496-508.

LaGasse LL, Neal AR, Lester BM (2005) Assessment of infant cry: Acoustic cry analysis and parental perception. In, pp 83-93.

Lehrman DS (1953) A critique of Konrad Lorenz's theory of instinctive behaviour. Quarterly Review of Biology 28:337-363.

Lorenz KZ (1950) The comparative method in studying innate behavior patterns. In, pp 221-268.

MacKay-Lyons M (2002) Central pattern generation of locomotion: a review of the evidence. Physical therapy 82:69-83.

Marques HG, Imtiaz F, Iida F (2013) Self-organization of reflexive behavior from spontaneous motor activity. Biological cybernetics 107:25-37.

Matsumoto D, Willingham B (2009) Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. Journal of personality and social psychology 96:1-10.

Mennella JA, Jagnow CP, Beauchamp GK (2001) Prenatal and Postnatal Flavor Learning by Human Infants. Pediatrics 107:e88-e88.

Merker H, Sames K, Csato W, Heger W, Neubert D (1988) The embryology of Callithrix jacchus. Non-human primates: developmental biology and toxicology Ueberreuter Wissenschaft, Wien:219-242.

Moon C, Cooper RP, Fifer WP (1993) 2-day-olds prefer their native language. Infant Behavior and Development 16:495-500.

Mori H (2012) Is the developmental order of fetal behaviors self-organized in an uterine environment? 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL):1-2.

Mori H, Kuniyoshi Y (2010) A human fetus development simulation: Self-organization of behaviors through tactile sensation. 2010 IEEE 9th International Conference on Development and Learning:82-87.

Narayanan CH, Fox MW, Hamburger V (1971) Prenatal development of spontaneous and evoked activity in the rat (Rattus norvegicus albinus). Behavior 40:100-134.

Newman JD (1985) The infant cry of primates. An evolutionary perspective. In: (Lester BM, Zachariah Boukydis CF, eds), pp 307-323. Boston, MA: Springer US.

Newman JD (2007) Neural circuits underlying crying and cry responding in mammals. In, pp 155-165.

Nowlan N (2015) Biomechanics of foetal movement. European cells & materials 29:1.

Nozaki S, Iriki A, Nakamura Y (1986) Localization of central rhythm generator involved in cortically induced rhythmical masticatory jaw-opening movement in the guinea pig. Journal of neurophysiology 55:806-825.

Oppenheim RW (1974) The ontogeny of behavior in the chick embryo. In, pp 133-172.

Pfeifer R, Iida F, Lungarella M (2014) Cognition from the bottom up: on biological inspiration, body morphology, and soft materials. Trends in Cognitive Sciences 18:404-413.

Pillai M, James D (1990) Are the behavioural states of the newborn comparable to those of the fetus? Early Human Development 22:39-49.

Piontelli A (2010) Development of normal fetal movements. Milan: Springer Verlag Italia doi 10:978-988.

Prechtl H, Fargel J, Weinmann H, Bakker H (1979) Postures, Motility and Respiration of Low-risk Pre-term Infants. Developmental Medicine & Child Neurology 21:3-27.

Prechtl HFR (1985) Ultrasound studies of human fetal behaviour. Early Human Development 12:91-98.

Prechtl HFR (2001) Prenatal and early postnatal development of human motor behaviour. Handbook of Brain and Behaviour in Human Development:415-427.

Provine RR (1980) Development of between-limb movement synchronization in the chick embryo. Developmental psychobiology 13:151-163.

Rallu M, Corbin JG, Fishell G (2002) Parsing the prosencephalon. Nature Reviews Neuroscience 3:943-951.

Reissland N, Francis B, Mason J, Lincoln K (2011) Do facial expressions develop before birth? Plos One 6.

Robertson SS (1987) Human cyclic motility: fetal-newborn continuities and newborn state differences. Developmental psychobiology 20:425-442.

Robinson SR, Smotherman WP (1992) Fundamental motor patterns of the mammalian fetus. Journal of neurobiology 23:1574-1600.

Robinson SR, Kleven GA (2005) Learning to move before birth. In, pp 131-175.

Robinson SR, Blumberg MS, Lane MS, Kreber LA (2000) Spontaneous motor activity in fetal and infant rats is organized into discrete multilimb bouts. Behavioral neuroscience 114:328-336.

Roodenburg PJ, Wladimiroff JW, van Es A, Prechtl HFR (1991) Classification and quantitative aspects of fetal movements during the second half of normal pregnancy. Early Human Development 25:19-35.

Ryder GH (1943) Vagitus Uterinus. American Journal of Obstetrics & Gynecology 46:867-872.

Sai FZ (2005) The role of the mother's voice in developing mother's face preference: Evidence for intermodal perception at birth. Infant and Child Development 14:29-50.

Sakamoto T, Nonaka S, Katada A (1997) Control of respiratory muscles during speech and vocalization. Neural Control of the Respiratory Muscles:249-258.

Schaal B, Rouby C (1990) Development of the chemical senses exogenous prenatal influences postnatal consequences. In: XXTH National Neonatal Conference, pp 182-201. Paris, France.

Sharp AA, Ma E, Bekoff A (1999) Developmental changes in leg coordination of the chick at embryonic days 9, 11, and 13: uncoupling of ankle movements. Journal of neurophysiology 82:2406-2414.

Shuleikina KV, Gladkovich NG (1965) Features and level of structural organization of the first reflexes in the human fetus. Bulletin of Experimental Biology and Medicine 59:226-229.

Sirkin DW, Schallert T, Teitelbaum P (1980) Involvement of the pontine reticular formation in head movements and labyrinthine righting in the rat. Experimental Neurology 69:435-457.

Smotherman M, Kobayasi K, Ma J, Zhang S, Metzner W (2006) A mechanism for vocal-respiratory coupling in the mammalian parabrachial nucleus. The Journal of neuroscience 26:4860-4869.

Smotherman WP, Robinson SR (1985) The rat fetus in its environment: Behavioral adjustments to novel, familiar, aversive, and conditioned stimuli presented in utero. Behavioral Neuroscience 99:521-530.

Smotherman WP, Robinson SR (1988) The uterus as environment. The ecology of fetal behaviour. Handbook of Behavioral Neurobiology 9:149-196.

Sorokin VN, Gay T, Ewan WG (1980) Some biomechanical correlates of jaw movements. The Journal of the Acoustical Society of America 68:S32-S32.

St James-Roberts I, Menon-Johansson P (1999) Predicting infant crying from fetal movement data: an exploratory study. Early Human Development 54:55-62.

Takahashi Daniel Y, Narayanan Darshana Z, Ghazanfar Asif A (2013) Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. In: Current Biology.

Takahashi DY, Fenley AR, Teramoto Y, Narayanan DZ, Borjon JI, Holmes P, Ghazanfar AA The developmental dynamics of marmoset monkey vocal production. In press.

Tardif SD, Mansfield KG, Ratnam R, Ross CN, Ziegler TE (2011) The marmoset as a model of aging and age-related diseases. ILAR Journal 52:54-65.

Wichterle H, Turnbull DH, Nery S, Fishell G, Alvarez-Buylla A (2001) In utero fate mapping reveals distinct migratory pathways and fates of neurons born in the mammalian basal forebrain. Development 128:3759-3771.

Windle WF (1944) Genesis of Somatic Motor Function in Mammalian Embryos: A Synthesizing Article. Physiological Zoology 17:247-260.

Winter P HPPD, Schott D (1973) Ontogeny of Squirrel Monkey Calls under Normal Conditions and under Acoustic Isolation. Behaviour 47:230-239.

Yamada Y, Kuniyoshi Y (2012) Embodiment guides motor and spinal circuit development in vertebrate embryo and fetus. 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL):1-6.

Yamada Y, Mori H, Kuniyoshi Y (2010) A Fetus and Infant Developmental Scenario: Self-organization of Goal-directed Behaviors Based on Sensory Constraints. Proceedings of the Tenth International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems.

Zeskind PS, Lester BM (2001) Analysis of infant crying. In: (L. T. Singer PSZ, ed), pp 149-166.