

Acoustic analysis of vocal development in a New World primate, the common marmoset (*Callithrix jacchus*)^{a)}

Ashley L. Pistorio, Brett Vintch, and Xiaoqin Wang^{b)}

Laboratory of Auditory Neurophysiology, Department of Biomedical Engineering,
Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

(Received 11 January 2006; revised 9 June 2006; accepted 21 June 2006)

In contrast to humans and songbirds, there is limited evidence of vocal learning in nonhuman primates. While previous studies suggested that primate vocalizations exhibit developmental changes, detailed analyses of the extent and time course of such changes across a species' vocal repertoire remain limited. In a highly vocal primate, the common marmoset (*Callithrix jacchus*), we studied developmental changes in the acoustic structure of species-specific communication sounds produced in a social setting. We performed detailed acoustic analyses of the spectral and temporal characteristics of marmoset vocalizations during development, comparing differences between genders and twin pairs, as well as with vocalizations from adult marmosets residing in the same colony. Our analyses revealed significant changes in spectral and temporal features as well as variability of particular call types over time. Infant and juvenile vocalizations changed progressively toward the vocalizations produced by adult marmosets. Call types observed early in development that were unique to infants disappeared gradually with age, while vocal exchanges with conspecifics emerged. Our observations clearly indicate that marmoset vocalizations undergo both qualitative and quantitative postnatal changes, establishing the basis for further studies to delineate contributions from maturation of the vocal apparatus and behavioral experience. © 2006 Acoustical Society of America. [DOI: 10.1121/1.2225899]

PACS number(s): 43.80.Ka [JAS]

Pages: 1655–1670

I. INTRODUCTION

It is commonly held that relatively little learning is involved in primate vocal communication, particularly with regard to vocal production (Seyfarth and Cheney, 1997). Past studies on the ontogeny of vocalizations in primates showed few significant changes in the acoustic structure of vocalizations (Seyfarth and Cheney, 1986; Seyfarth and Cheney, 1997; Biben and Bernhards, 1995; Fischer, 2003; Inoue, 1988; Kalin and Shelton, 1998; Masataka, 1993; McCowan *et al.*, 2001; Merker and Cox, 1999; Newman, 1995; Newman and Symmes, 1974; Owren *et al.*, 1992; Owren *et al.*, 1993; Snowdon, 1997; Snowdon *et al.*, 1997; Winter, 1969; Winter *et al.*, 1973). One study showed morphological changes in the acoustic parameters of developing squirrel monkey vocalizations, but these changes were attributed strictly to maturation (Hammerschmidt *et al.*, 2001). Many of the studies investigating changes in acoustic morphology during vocal development did not employ extensive quantitative analysis methods due to the limitations of audio recording quality in the field and analytical techniques. A careful characterization of ontogenetic changes in the acoustic structure of vocalizations is a necessary first step for future experiments intended to delineate physical maturation from potential vocal production learning in nonhuman primates

(Egnor and Hauser, 2004). Collecting and quantitatively analyzing a large sample set of vocalizations from multiple animals of a single species through the early stages of physical and social development is the best way to achieve this goal.

In the present study, we investigated the acoustic changes that occurred in vocalizations during the ontogenetic development of an arboreal New World Primate species native to the Atlantic coastal forests of Brazil, the common marmoset (*Callithrix jacchus*). Marmosets have a broad vocal repertoire and live in small groups, where only the dominant male and female breed but other group members participate in raising infants (Stevenson and Poole, 1976). Observations in the field suggest that common marmosets use vocalizations to maintain group cohesion, to announce food, to warn of predators, to announce territory to other groups, and to locate their own group members over distances (Epplé, 1968; Moody and Menzel, 1976). In captivity, marmosets maintain an extensive repertoire of vocal communication sounds, some of which have been the subject of behavioral studies (Norcross and Newman, 1993, 1997; Schrader and Todt, 1993; Miller and Wang, 2006). Most of these studies investigated the Phee call because they are produced in a clear social context (isolation) and are easily elicited. Behavioral observations show that typically upon hearing a conspecific produce a Phee, marmosets will emit a Phee in response; this type of behavior is known as antiphonal calling. In their study, Miller and Wang (2006) showed that antiphonal calling in the common marmoset is dynamic; changes to the sensory input (i.e., the presence of a conspecific, identity and sex of the caller, and whether the occluded

^{a)}Portion of this work presented in "Ontogenetic Changes in Vocal Production and Usage in Common Marmoset Monkeys," Society for Neuroscience Abstract No. 541.1(2001).

^{b)}Author to whom correspondence should be addressed. Tel. (410) 614-4547; electronic mail: xiaoqin.wang@jhu.edu

individual participated in reciprocal antiphonal exchanges) result in modulations of vocal output. Other types of marmoset vocalizations that are produced in more ambiguous social contexts than Phee calls have not been systematically examined in any type of study.

The present study represents a comprehensive analysis of vocal development in marmosets. Quantitatively documenting developmental changes in the acoustics of marmoset vocalizations is the first step in creating a model for testing the influence of physical maturation, sensory experience, and learning in vocal development of a nonhuman primate. In this report, we examine the acoustic morphology of marmoset vocalizations produced during development. We accomplished this by taking advantage of the common marmoset's high vocal output rates in captivity to conduct vocal recordings in the social setting of a large breeding colony. Following these recordings, we then used advanced signal acquisition and processing techniques to analyze a large set of vocalization samples collected at multiple time points during development from single infants and twin pairs of different social groups within the colony. Utilizing sensitive acoustic analysis tools, we quantified changes in the acoustics of vocalizations and showed the course of normal vocal development in the common marmoset. Our results showed that marmoset vocalizations exhibit developmental changes in acoustic morphology.

II. METHODS

A. Subjects

This study included data from nine common marmosets (*C. jacchus*), born into a captive colony of approximately 50 of the same species. Eight were part of twin pairs and one was a single infant; subjects were followed from age 3 weeks to age 25 weeks, the period over which the most noticeable vocal production changes take place. The colony, which included family and social groups ranging in ages from infancy to ~12 years, was kept in spacious cages furnished with natural hardwood branches, nesting boxes, hammocks, and various other physical enrichment items.

B. Vocalization recording

We conducted all recordings without an experimenter in the room and video documentation was occasionally made for sessions with one twin pair. Infants were separated from their family group for brief intervals of 20 min to 2 h, depending on age: at 3–4 weeks of age, the time interval was 20 min; at 5–7 weeks, it was 40 min; at 8–12 weeks, it was 1 h; at 11+ weeks, it was 2 h. They were placed in a custom-made recording cage surrounded by 3 in. acoustic absorption foam (Sonex, Illbruck). Each infant was placed in the recording cage individually, to ensure clean and classifiable vocalization samples. Recordings were taken using two microphones (AKG C1000S), one facing the infant and the other facing its family group. Microphone output signals were amplified (Symetrix SX202) and then recorded using a two-channel professional digital audio tape (DAT) recorder (Panasonic SV-3700 or TASCAM DA-40) at a sampling rate of 48 kHz.

C. Data analysis

Experimenters screened all vocalizations recorded on DAT tapes via both audio and visual display of real-time spectrogram using a PC equipped with RTS software (Engineering Design, CA). Segments of acoustic data containing vocalizations captured from DAT tapes were transferred onto a PC hard drive and analyzed using custom software written in MATLAB (MathWorks, MA) programming language. We used three levels of analysis with this custom software to process the vocalization data in increasing detail. Statistical tests were then used to determine the significance of changes in the vocalization acoustics of individual animals during development, and contributions of gender and twin pairing in these changes.

1. Digitizing and categorizing vocal signals

The first level of analysis utilized custom software to automatically screen through the saved sound files and extract vocalizations from the animal of interest based on specified parameters for signal and noise levels. Each extracted vocalization was manually examined to assure it is complete and free of noise. Vocalizations which passed quality criteria were then classified by acoustic features into adultlike major call types (Twitter, Trill, Phee, Trillphee) and additional call types (Egg, Ock, Tsik, and Compound Calls) (Epplé, 1968; Agamaite, 1997). The three most commonly produced infant marmoset vocalizations that were unclassifiable as adultlike call types were noted in this study. We also observed various vocalizations which did not match the criteria as adultlike or the three major infant-specific calls. These vocalizations were variable in their acoustic structures and were not analyzed further.

2. Computing acoustic parameters

Acoustic features in each of the four major call types were quantified based on those previously analyzed in the adult marmoset in our laboratory (Agamaite, 1997), with modifications to detect novel features only observed in infants. Parameters analyzed included spectral, temporal, and amplitude characteristics (see Table I for a list of parameters).

3. Statistical analysis of vocalization features

Statistical analysis of vocalization features was carried out by call type, age, gender, and individuals. Trends of group data were calculated by taking average parameter values across animals for each week. Two variations of this calculation were made. For the individual mean (IM), the average was calculated by first taking the mean of each animal. Using the individual mean data, slopes of change over time for each animal were computed for each parameter and for each call type using the method of least squares. These individual means were also averaged to obtain the group mean, where each monkey factors equally. For the population mean (PM), the average was calculated from all calls recorded from every animal. No weighting is given to offset the number of calls produced by each animal.

TABLE I. Vocalization parameters. Parameter names are listed in the left column, with the corresponding parameter number for each of the four major call types (Phee: 1–30, Trill: 1–36, Trillphee: 1–36, and Twitter: 1–28), followed by a description of the parameter, in subsequent columns. Parameters with no corresponding number for any of the vocalizations indicate that parameter does not apply to the vocalization type.

Parameter	Phee	Trill	Trillphee	Twitter	Additional Description
Duration	1	1	1	1	Total duration of a call
Time _{AmpHi}	2	2	2	2	Time of the absolute high amplitude for the entire call
Front Amp _{Mean}	3	3	3	3	Relative mean amplitude, first phrase for the Twitter
Front Freq _{Dom}	4	4	4	4	Dominant frequency, first phrase for the Twitter
Front Freq _{QLo}	5	5	5	5	Quarter-level low frequency, first phrase for the Twitter
Front Freq _{QHi}	6	6	6	6	Quarter-level high frequency, first phrase for the Twitter
Front Freq _{Min}	7	7	7	7	Absolute minimum frequency, first phrase for the Twitter
Front Freq _{Mean}	8	8	8	8	Absolute mean frequency, first phrase for the Twitter
Front Freq _{Max}	9	9	9	9	Absolute maximum frequency, first phrase for the Twitter
Middle Amp _{Mean}	10	10	10	10	Relative mean amplitude, middle phrase for the Twitter
Middle Freq _{Dom}	11	11	11	11	Dominant frequency, middle phrase for the Twitter
Middle Freq _{QLo}	12	12	12	12	Absolute minimum frequency, middle phrase for the Twitter
Middle Freq _{QHi}	13	13	13	13	Quarter-level high frequency, middle phrase for the Twitter
Middle Freq _{Min}	14	14	14	14	Absolute minimum frequency, middle phrase for the Twitter
Middle Freq _{Mean}	15	15	15	15	Absolute mean frequency, middle phrase for the Twitter
Middle Freq _{Max}	16	16	16	16	Absolute maximum frequency, middle phrase for the Twitter
End Amp _{Mean}	17	17	17	17	Relative mean amplitude, last phrase for the Twitter
End Freq _{Dom}	18	18	18	18	Dominant frequency, last phrase for the Twitter
End Freq _{QLo}	19	19	19	19	Absolute minimum frequency, last phrase for the Twitter
End Freq _{QHi}	20	20	20	20	Quarter-level high frequency, last phrase for the Twitter
End Freq _{Min}	21	21	21	21	Absolute minimum frequency, last phrase for the Twitter
End Freq _{Mean}	22	22	22	22	Absolute mean frequency, last phrase for the Twitter
End Freq _{Max}	23	23	23	23	Absolute maximum frequency, last phrase for the Twitter
Absolute Freq _{Max}	24	24	24		Maximum frequency over the entire call
Time Abs Freq _{Max}	25	25	25		Time into the call of the maximum frequency
Absolute Freq _{Min}	26	26	26		Minimum frequency over the entire call
Time Abs Freq _{Min}	27	27	27		Time into the call of the minimum frequency
Frequency _{start}	28	28	28		First measurable frequency in a call signal
Frequency _{End}	29	29	29		Last measurable frequency in a call signal
Time _{Transition}	30	30	30		Time of transition from Trilling to Pheeing (value=0 for Phee)
Time _{Period}		31	31		Time of a trilling cycle period
Δ Freq _{Max}		32	32		Maximum frequency change in one trilling cycle
Time Δ Freq _{Max}		33	33		Time it takes for the maximum frequency change in a cycle
Δ Freq _{Min}		34	34		Minimum frequency change in one trilling cycle
Time Δ Freq _{Min}		35	35		Time it takes for the minimum frequency change in a cycle
Δ Freq _{Mean}		36	36		Mean frequency change in one trilling cycle
IPI				24	Average inter-phrase interval for all phrases of a Twitter call
Number Phrases				25	Total number of phrases in a Twitter call
Time _{Sweep} Phr _{First}				26	Time for the first phrase frequency sweep
Time _{Sweep} Phr _{Middle}				27	Time for the middle phrase frequency sweep
Time _{Sweep} Phr _{Last}				28	Time for the last phrase frequency sweep

a. Distance measures: A useful comparison of vocalizations between infant and adult groups, as well as between individual animals at a specific age, can be made using distance measurements in a multidimensional space, where each dimension represents a particular vocalization feature. We compute distances between comparison groups using multidimensional Euclidean distance. In a two- or three-dimensional space, the Euclidean distance is the geometric distance between two points. Because the Euclidean distance can be biased by differences in scale of each of the dimensions used to compute it, it is good practice to transform the dimensions so they have similar scales. Here we measured the Euclidean distance between two groups of data using parameter values normalized by the variance along each dimension, as follows:

$$D(x,y) = \sqrt{\sum_{i=1}^N \left(\frac{x_i - y_i}{\sigma_i} \right)^2}, \quad (1)$$

where D is the N -dimensional Euclidean distance between data group X and group Y ; x_i and y_i are parameter values on the i th dimension for each data group, respectively; σ_i is the variance of the parameter values along the i th dimension.

b. Differences between gender, sibling pair, and age groups: An analysis of nonindividual factors in data trends is accomplished using statistical covariance tests. Animals were either grouped by gender or sibling pairs and their vocalization trends (both by absolute parameter values and by distance from adulthood) were compared with analysis of

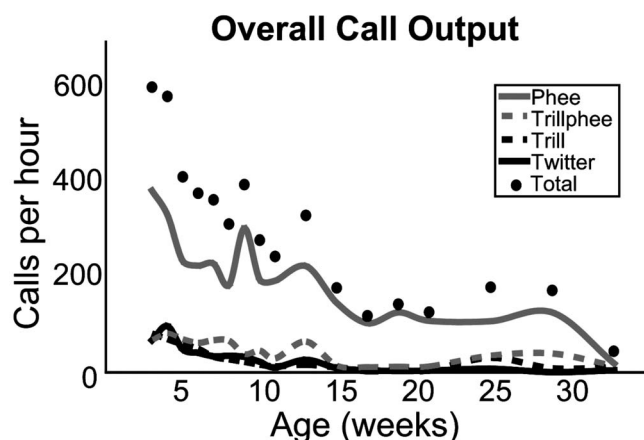


FIG. 1. Population average of call output (in calls per hour) as a function of age (in weeks) for all infant marmosets used in this study. Both the total number of calls (black circles) and the number of calls for each of the four major call types are shown.

covariance (ANCOVA). Using a confidence level of 0.05, the likelihood that vocalization trends were similar between gender groups and sibling pairs was assessed along multiple dimensions.

In evaluating the significance of differences between mean values in call parameters, we also grouped vocalization samples by age and gender using individual parameter distributions. Significance in mean difference between each set of groups was assessed using a two-tailed t-test at both $\alpha=0.05$ and $\alpha=0.01$.

Principal component analysis (PCA) was used to determine the call parameters that most explain variance between age groups. Vocalizations from 5 weeks and 25 weeks were used to simplify the process while maintaining broad age separation. Weighting vectors were found that described a transform to project the data into an orthogonal basis set describing maximum variance. The highest weighted parameters in the principal component vector explained the most amount of variance.

c. Discriminability of vocalizations between individual animals: Discriminability of vocalizations between individual animals was determined using a discrimination algorithm. The vocalizations were first reduced to a two-dimensional representation using the first two PCA

components. Then, for all four major call types, the mean and distribution of principal components for each animal's set of vocalizations were calculated. Using this information, each individual vocalization's distance is measured from the animal means, normalized by standard deviation, and a minimum distance is found. A vocalization is considered correctly classified if the smallest distance belongs to the mean of animal from whom the vocalization was recorded.

III. RESULTS

A. Qualitative description of general observations

1. Overall call output

Production of all call types decreased in occurrence with advancing age. The overall call output averaged over the population of infants from this study is plotted in Fig. 1. The Phee call is the most commonly-produced call, while the other three major call types, often used in close (Trill, Trillphee) or distant (Twitter) contact situations, were infrequently produced in older animals physically isolated from their family group but with visual contact maintained.

2. Vocal exchanges between infant subjects and other conspecifics

For vocalizing infants aged 3–8 weeks, there were no discernable vocal responses from adults. Infant subjects often vocalized shortly after (≤ 1 s) vocalizations by other marmosets with the same call type. The infant subjects rarely decreased or ceased vocal output when other animals in the colony alarm-called or became suddenly silent. Several examples of infant responses to calls made by adult conspecifics are shown in Fig. 2. Infant responses occurred within an average of 600 ms after the offset of the conspecific call in the examples shown in Fig. 2(a) that include all four major call types. Figure 2(b) shows examples of infants vocalizing Phee strings (which sometimes began with a Trillphee) in response to Phees made by adult animals in the colony.

As the infants grew into juveniles, and then adults, there gradually emerged more directed responses based on conspecific vocalizations, and juveniles became silent in response to colony alarms. Responses of conspecifics to the subjects

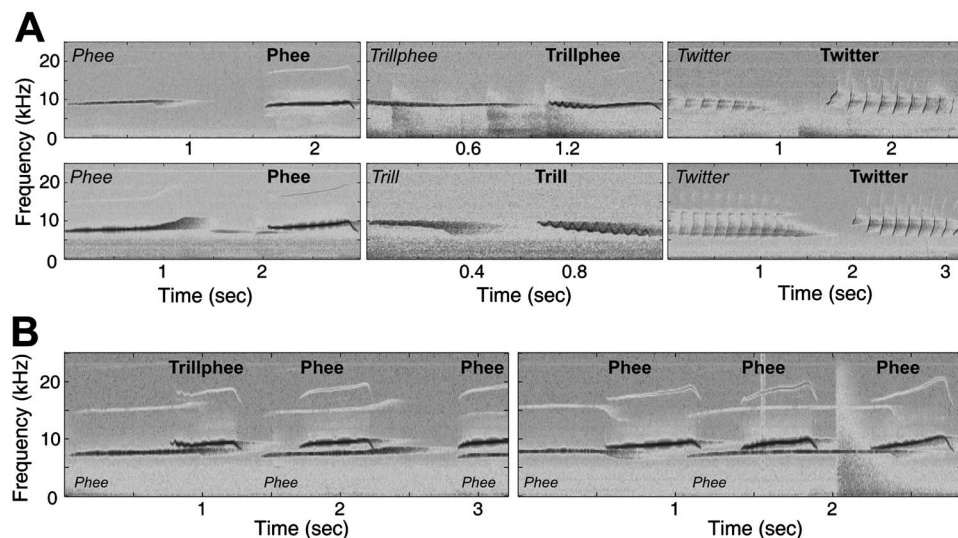


FIG. 2. Examples of 3–4 week old infant vocal responses to adult conspecifics. (A) Spectrograms, labeled for call type (*italic*: adult vocalization; **bold**: infant response) are shown for the four major call types. Each plot shows one instance of adult-infant vocalizations. (B) Two instances of Phee strings made by an infant in response to Phee calls from an adult are shown. Note the higher frequency of infant calls as compared to adult calls in both plots and the timing of infant calls (lagged behind adult calls).

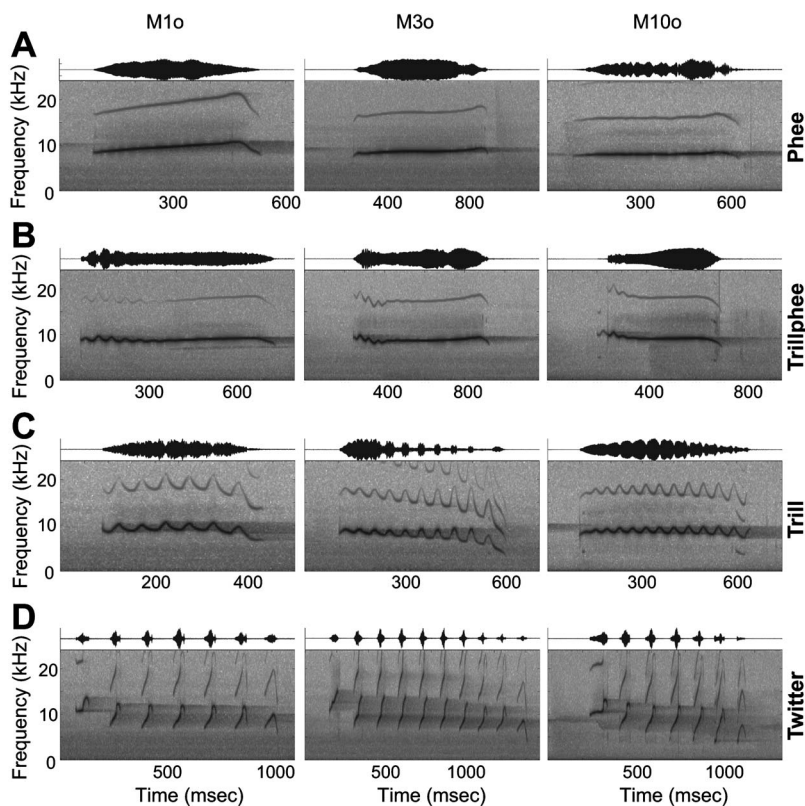


FIG. 3. Spectrogram and amplitude waveform exemplars are shown for the four major call types: (A) Phee; (B) Trillphee; (C) Trill; (D) Twitter, from three different (unrelated) infants (M10o, M30o, M100o) at 3–5 weeks of age.

were difficult to determine in this study, because we did not perform behavioral scoring and the vocal output of juveniles included comparatively fewer calls than that of infants.

B. Call types and acoustic characteristics in marmoset infants

Most common marmoset calls are tonal and can be grouped into call types that remain consistent across individuals. To analyze the vocal development of each call type in infant marmosets, we used parameters defined for the normal adult calls of the same type and quantified changes in each parameter across different ages in the same animals. Additional parameters were used to accommodate acoustic features unique to infant calls.

1. Adultlike call types and novel features in infants

Infant marmosets begin vocalizing shortly after birth and produce some call types which are acoustically similar to the call types produced by adults. Figure 3 shows examples of the four major call types (Phee, Trillphee, Trill, and Twitter), also produced by adult animals, made by three unrelated infant marmosets at 3–5 weeks of age. Both spectrograms and amplitude waveforms are shown. The Phee is a tonelike long call; the Trillphee is similar to a Phee, but with sinusoidal-like amplitude modulation (AM) and frequency modulation (FM) during the first portion of the call. Trill calls are defined by sinusoidal-like AM and FM throughout the entire call. Twitters consist of multiple phrases of upward FM sweeps, some of which ended by brief downward FM sweeps. The first and last phrases of twitters are highly variable in structure.

Progressive changes over time for the four major call types are shown by representative examples in Fig. 4. For

example, the Phee and Trillphee calls become longer in duration and lower in frequency over time and the FM depth decreases for the Trillphee and Trill calls. Several vocalization features are seen often in infants, but are apparently not produced in adults. One such feature is the addition of spectral components outside the predominant harmonic structure, presumably related to immaturity of the vocal apparatus [Fig. 4(a), 4-week Phee example, closed arrow]. Another infant-specific feature is the “twitter-hook,” a short downward FM sweep at the end of each upward FM sweep in each phrase of a Twitter call [Fig. 4(d), open arrow]; the hooking fades as an animal matures. All infant marmosets in this study exhibited this feature at the earliest time of observation (3–5 weeks) but varied in the extent of the downward FM sweep, its temporal span, and the age at which it faded.

We also observed other, less frequently produced, adultlike call types in infants. Figure 5(a) shows examples of Peeps, similar to a Phee in frequency and structure, but significantly shorter in duration. Figure 5(b) shows a commonly-heard alarm call, the Tsik, and one example of a Tsik followed by an Egg (short-duration harmonic stack, indicated by an arrow), a configuration seen in adults as well. Figure 5(c) shows examples of compound calls, occasionally observed in adults, where the infant uses multiple call types strung together with little or no intersyllable interval.

2. Infant-specific calls

While marmosets produced the four major adultlike call types at all stages of development, three dominant patterns of vocalization emerged in infant marmoset vocalizations: the Cry, Compound Cry, and call strings or “Babbling.” The Cry is an extremely broadband, long-duration call [Fig. 6(a)]. In-

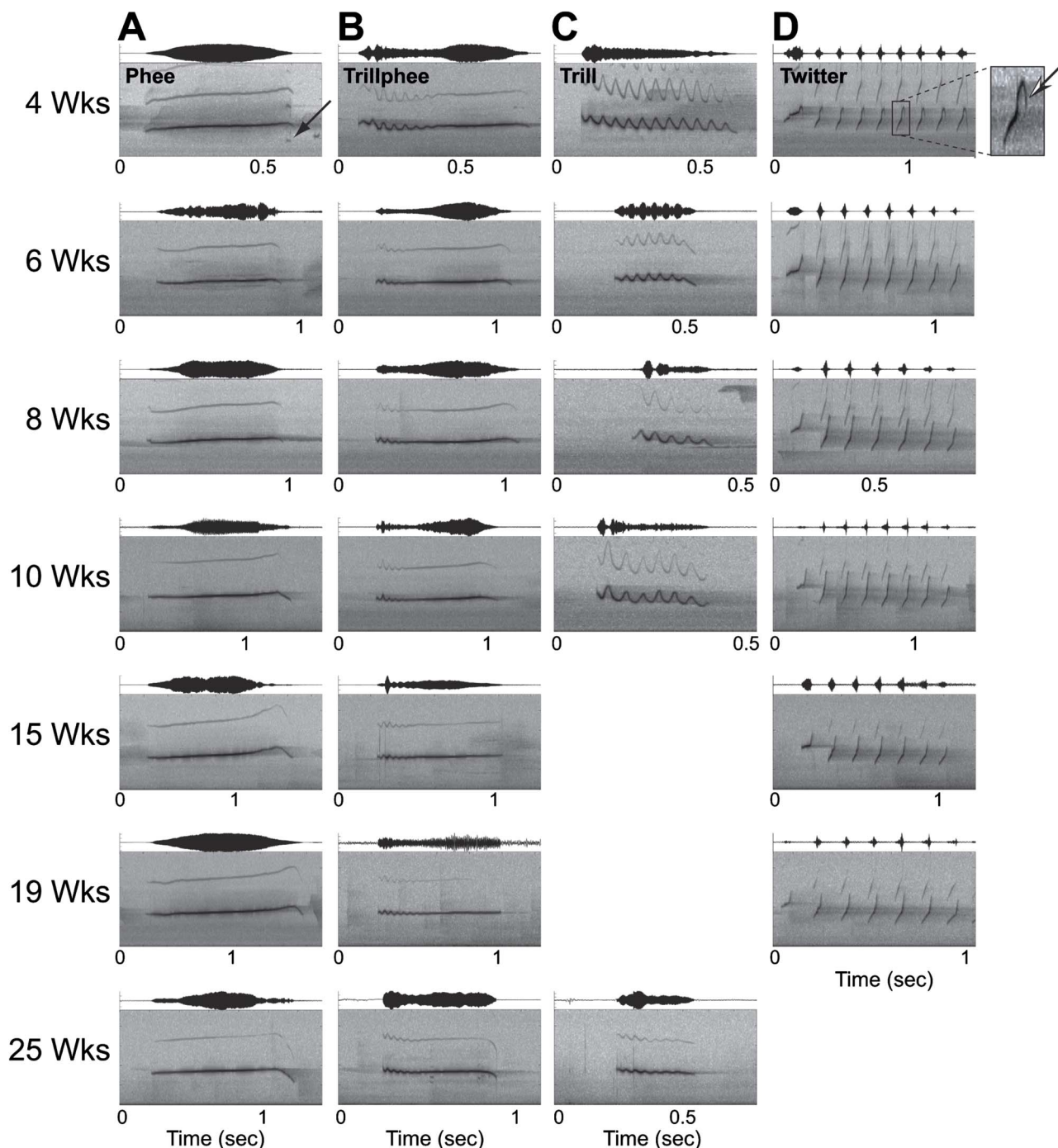


FIG. 4. Exemplars showing the progression of the four major call types as one individual marmoset (m10o) grew (4, 6, 8, 10, 15, 19, and 25 weeks of age). (A) Phee; (B) Trillphee; (C) Trill; (D) Twitter. Closed arrow indicates the appearance of spectral components outside the predominant harmonic structure of the Phee call. Open arrow indicates the “twitter hook.”

fant marmosets often use the Cry in compound calls with other call types, as seen in Fig. 6(b). Another infant-specific vocal behavior is the production of continuous strings of multiple call types which can last for minutes. Figure 6(c) shows 38 s of one such string. These call strings are similar in nature to the “babbling” seen in pygmy marmosets by Snowden and Elowson (2001). At ages up to 6 or 7 weeks, the Cry, Compound Cry, and Babbling were the most common vocalizations uttered by marmosets. However, the Babbling abruptly disappeared, usually by the seventh week, and the use of the Cry and Compound Cry faded gradually and

disappeared entirely by 10–11 weeks in all animals in the study. These infant-specific call types or behaviors were not observed in a previous study on classification and acoustic quantification of adult marmoset call types from this same captive colony (Agamaite, 1997).

C. Quantification of changes in vocalization features over time

For each of the four major adult-like call types, we quantified changes in a set of spectral and temporal param-

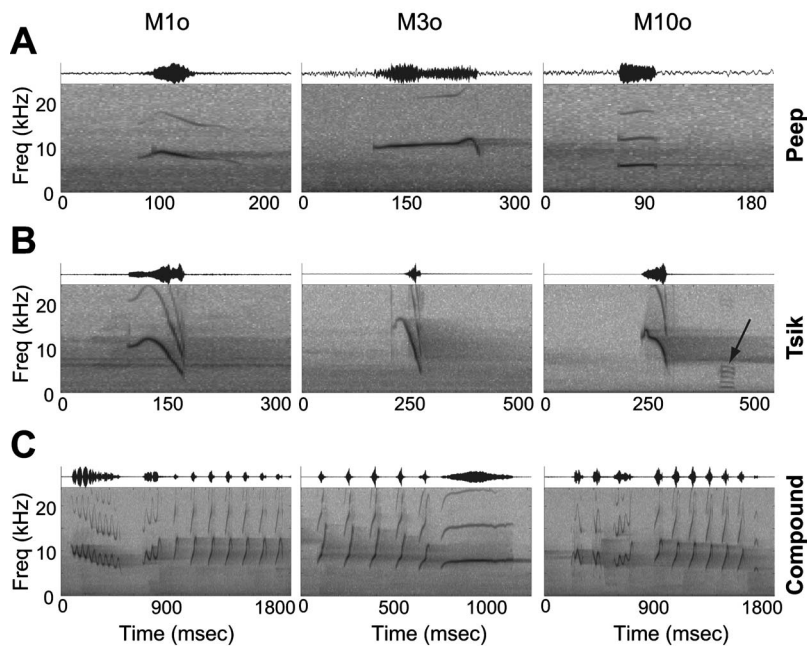


FIG. 5. Spectrogram and amplitude waveforms of other adult-like call types; (A) Peep; (B) Tsik; (C) Compound, produced by three different (unrelated) infants (M1o, M3o, M10o) at 3–5 weeks of age. The arrow [(B) third column] highlights an Egg call, a harmonic stack, occurring after the tsik.

eters (see Table I) as an animal aged. We then compared the differences between individuals, gender, and siblings in their call parameters at different ages, identifying which features become more dissimilar and which features become more stereotyped over time.

1. Individual subjects

Examples of the typical acoustic changes of vocalizations that occurred over time in the four major call types are shown in Fig. 7. Each line represents the least-squares fit of the individual mean (IM) for one vocalization parameter measured at discrete time points from one of the infant marmosets in this study. While some spectral and temporal parameters changed over time at similar rates (i.e., with similar slopes of lines in Fig. 7) for all sampled animals, other parameters showed different changes over time among different animals. We further quantified the rate of acoustic changes over time by computing the slope of least-squares fit for each vocalization parameter measured at different ages.

Figures 8 and 9 show the distribution of the slopes of the least-squares fit for each measured parameter of each call type for all individual animals using the IM method. These boxplots are representative of the rate of change in Hz per week for spectral parameters, seconds per week for temporal parameters, and relative dB per week for amplitude parameters. For each parameter, the area within the box corresponds to the middle 50% of the slopes for all animals, with the median indicated by the black center line. Outliers (≥ 1.5 times the range from the upper and lower quartile) are marked by the short lines outside of the whiskers. The rates of change for all measured spectral parameters of the four major call types are shown in Fig. 8. Slopes of the measured temporal parameters are plotted in Fig. 9(a) and slopes of the measured relative amplitude parameters are plotted in Fig. 9(b).

Phee [Figs. 7(a), 8(a), 9(Aa), and 9(Ba)]: Most spectral parameters for the Phee [Fig. 8(a)] showed trends for each

animal in the same general direction, but to varying degrees. Overall, across the population, most frequency parameters for the Phee decreased in their respective values. Temporal parameters for the Phee varied in magnitude of slope for different animals but all were in a positive direction except for one parameter [Fig. 9(Aa)]. A consistent trend across animals for amplitude parameters was that the beginning of the Phee became lower in relative amplitude [param. 3, Fig. 9(Bb)], while the end of the Phee increased in amplitude [param. 17, Fig. 9(Bb)] as the animals aged.

Trillphee [Figs. 7(b), 8(b), 9(Ab), and 9(Bb)]: Spectral parameters of Trillphee [Fig. 8(b)] are less variable than the Trill [Fig. 8(c)] but more so than the Phee [Fig. 8(a)]. Although Trillphee frequency parameters largely showed a downward slope, several animals increased frequency parameters as they aged (beginning of call: param. 4–9; middle of call: param. 11–16; end of call: param. 18–23). As in the Phee, the duration (param. 1) and time of maximum amplitude (param. 2) increased as animals grew older [Fig. 9(Ab)]. However, unlike the Phee, the time of high frequency (param. 25) decreased and the time of low frequency (param. 27) increased, on average, as animals matured [Fig. 9(Ab)]. Relative amplitude parameters for the Trillphee showed little change overall across individual animals, with the most consistent change being an increase in amplitude at the beginning of the call (param. 3) relative to the middle (param. 10) and end of the call (param. 17) [Fig. 9(Bb)].

Trill [Figs. 7(c), 8(c), 9(Ac), and 9(Bc)]: Frequency parameter changes in the Trill [Fig. 8(c)] for each section of the call were widely variable across individuals. There were outliers with high positive slope values for all nonmodulation-related spectral measures. On average, however, the slope showed little change or a decrease in these parameters as animals aged. The modulation frequency measures (param. 32, 34, and 36) also decreased, indicating a sinusoidal shape of consistently decreasing bandwidth as animals matured. Temporal parameters for the Trill [Fig. 9(Ac)] showed some

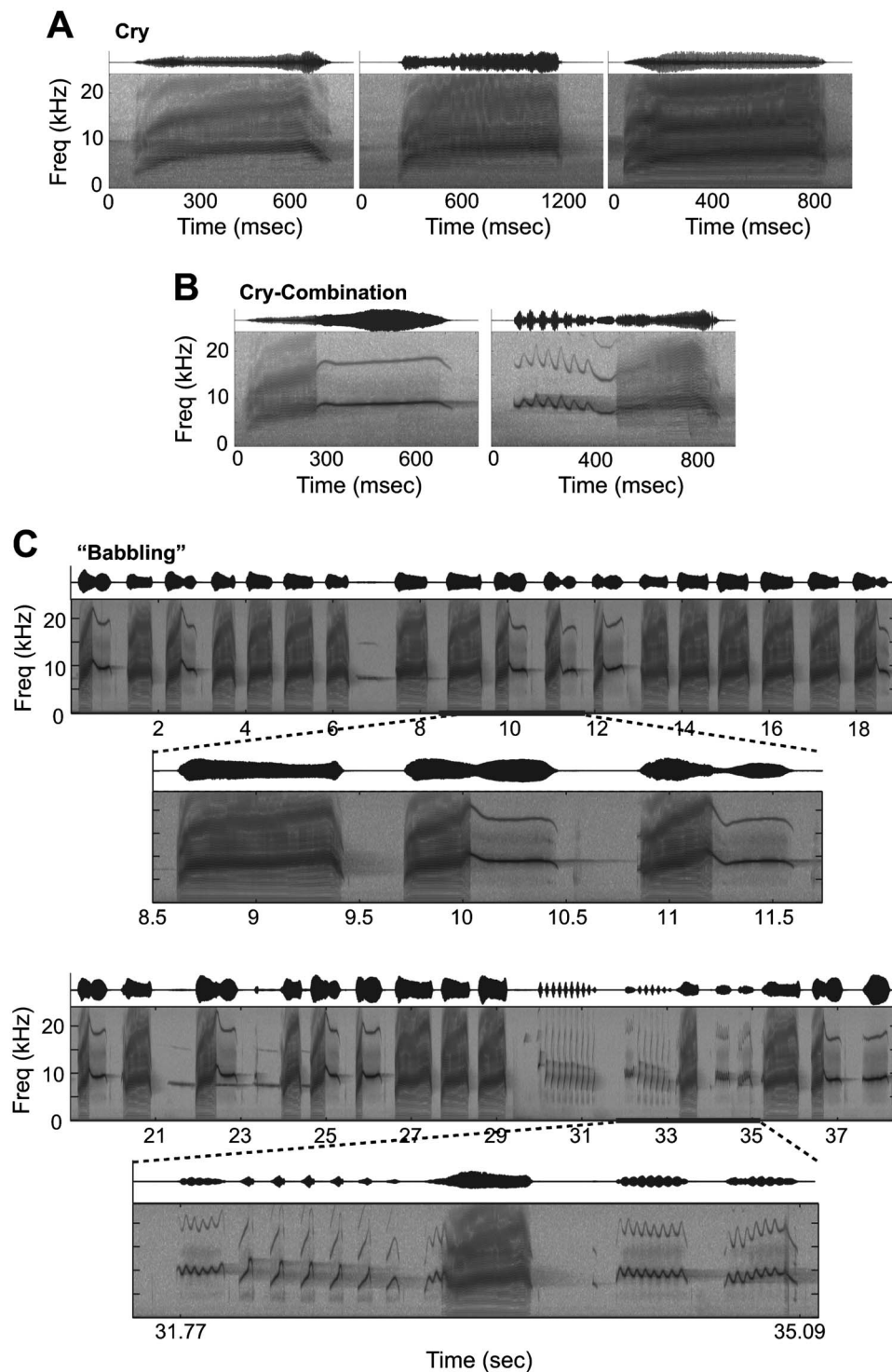


FIG. 6. Example call types produced exclusively by infants. (A) Cry calls; (B) Cry-combination calls; (C) "Babbling:" a 38-s continuous string of vocalizations broken down into two parts for ease of visualization. A small segment of each part is magnified to show detail.

changes similar to those seen in the Trillphee, but to a lesser extent. The range of slope change for two temporal parameters associated with the frequency modulation (param. 33 and 35) was significantly broader than that seen in the Trillphee. The average relative amplitude [Fig. 9(Bc)] of the middle portion of Trill (param. 10) decreased, although its range was highly variable, and the average amplitude of the end portion of the Trill (param. 17) increased.

Twitter [Figs. 7(d), 8(d), 9(Ad), and 9(Bd)]: The spectral

structure of the first phrase of each utterance of a Twitter call is extremely variable, contributing to the wider variation in slopes for parameters 3–9 [Fig. 8(d)]. The middle phrase, however, is more stable in comparison and frequency parameter slopes (param. 11–16) indicate an overall decrease in frequency over time for most. The Twitter duration increased for all but one animal, which showed a very small decrease, and the time of maximum amplitude had a positive slope for all animals [Fig. 9(Ad)]. For the Twitter call, most animals

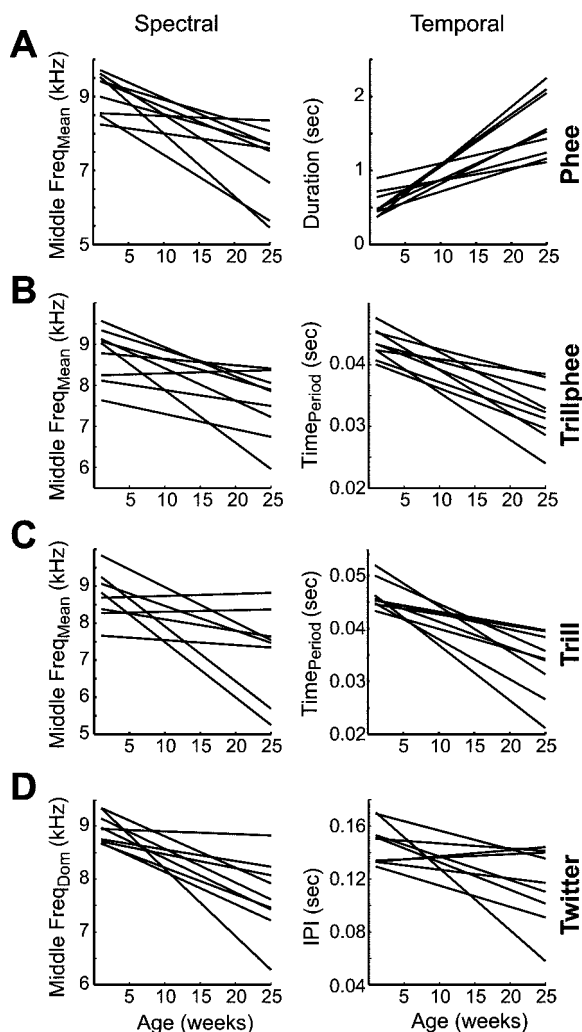


FIG. 7. Least-squares regression analysis of representative spectral (left) and temporal (right) parameters for the four major call types: (A) Phee; (B) Trillphee; (C) Trill; (D) Twitter, for all studied individuals. Parameter value is on the y axis and age (in weeks) is on the x axis. Each best-fit line corresponds to IM data from an individual animal as it aged.

showed a decrease in relative amplitude of the first and last phrases as they aged, while the middle phrase relative amplitude increased in all but one animal, [Fig. 9(Bd)].

2. Population data

For the entire population of tested animals, we show several example parameters of each of the four major call types in Fig. 10 to illustrate the general trends of individual acoustic features as an animal develops. The individual (IM) and population (PM) means are shown for comparison in each plot. Note that for many parameters shown, at early time points, the population and individual means have similar values (and individual mean error is small) but they become more different (and individual mean error becomes larger) as the animals get older. Population measures of parameter changes for the Trillphee [Fig. 10(b)] and Trill [Fig. 10(c)] calls show greater variations among individuals, especially as they became older, as compared to the Phee. The

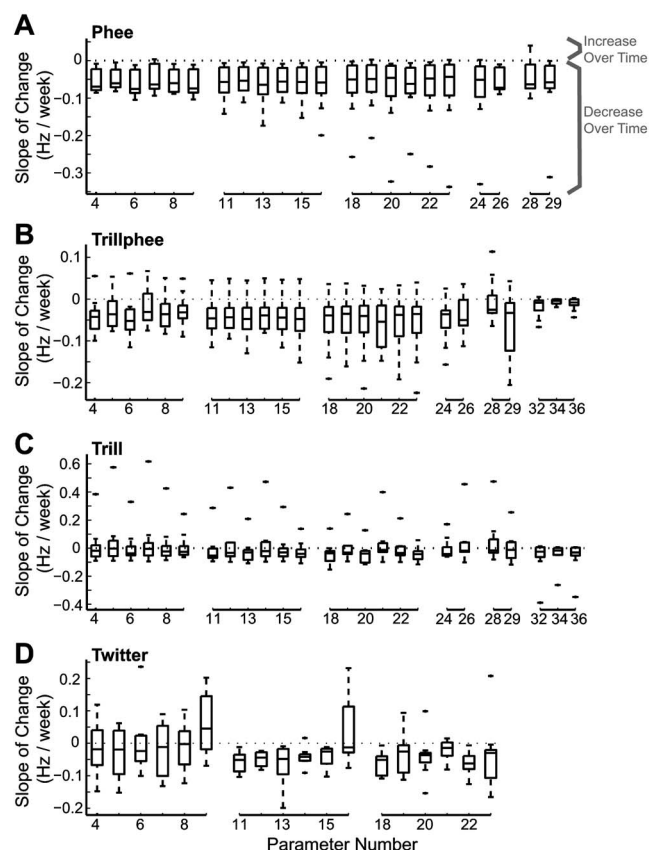


FIG. 8. Overall developmental trends of all measured spectral parameters for the four major call types: (A) Phee; (B) Trillphee; (C) Trill; (D) Twitter is analyzed by the slope of the least-squares regression line (see Fig. 7) of each parameter for each individual animal. The slope (indicating the magnitude and direction of the change in a parameter over time, in Hz/week) is on the y axis and the parameter number (see Table I) is on the x axis. The horizontal dotted line at $y=0$ indicates the dividing point where data above the line represent a positive slope and data points below the line indicate a negative slope, as shown on the right of plot (A). For each parameter, mean slope values from all studied individual animals are shown as boxplot. The median of all data points for each parameter is indicated by the black centerline within the box, and the first and third quartiles are the top and bottom edges of the box. Extreme values (≤ 1.5 times the range from the upper and lower quartile) are the ends of the lines extending from the box. Points ≥ 1.5 times the interquartile range (outliers) are plotted as individual points. Boxplots are subdivided into groups (as indicated by a break in the x axis) representing beginning, middle, and end of the call, as well as overall parameters (such as start/end frequency and min/mean/max frequency).

population parameters for the Twitter [Fig. 10(d)] became more variable as animals aged, as indicated by the increased error bars.

3. Comparison of major call types between infants and adults

Infant data at 5 weeks was compared with data previously obtained from the same captive colony in adult marmosets (Agamaite, 1997) for the same parameters across the four major call types. Examples of individual parameter distributions are shown in Fig. 11. In the call parameters where distributions differed between infants and adults, the individual mean and distributions of the measured parameters for infants always trended towards the adult mean and distribution as they aged.

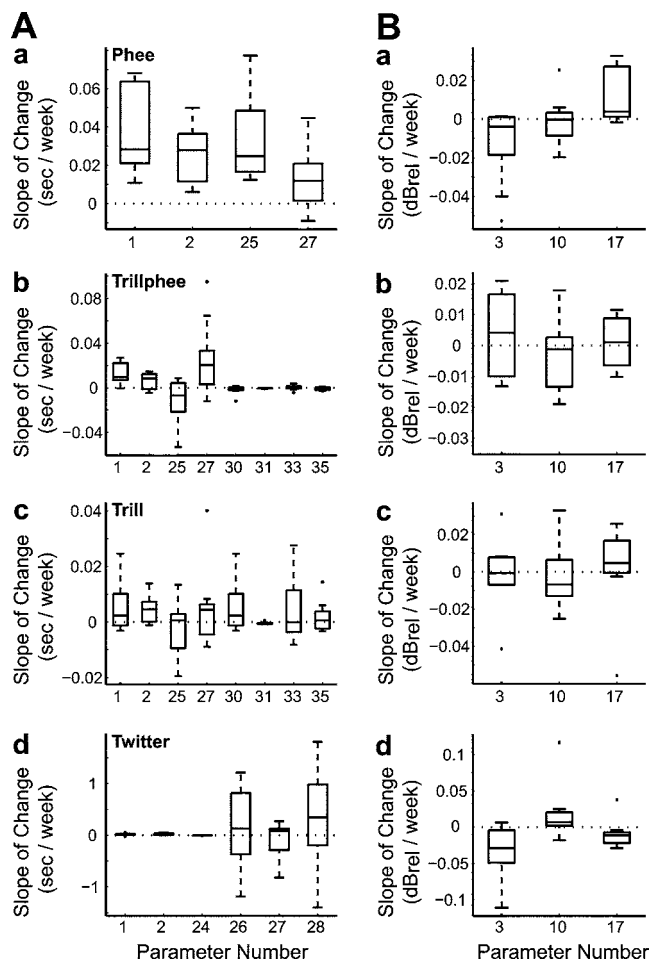


FIG. 9. Overall developmental trends of all measured temporal (left column) and all measured amplitude (right column) parameters for the four major call types (A) Phee; (B) Trillphee; (C) Trill; (D) Twitter. The format is the same as in Fig. 8, except there are no breaks in the x axis because each parameter number is labeled.

Euclidean distance [as defined in Eq. (1)] was calculated using the 18 spectral parameters common to all four major call types between the infants at different ages versus themselves at 25 weeks and between the same infants and adults from an earlier study (Agamaite, 1997). The results show that parameter changes are very similar between the group comparisons [Fig. 12(a)]. Generally, the Phee showed the largest and most rapid change in parameter measurements over time, while the Trillphee and Trill showed intermediate changes. The Twitter showed only a small, gradual change. The percent change for these comparisons, shown in Fig. 12(b), is also not significantly different between groups. These data indicate that the measured acoustic parameters of marmoset vocalizations stabilize by the time animals reach the age of about 25 weeks.

The standard deviation (SD) of call parameters between all studied infants is plotted against those of adults studied earlier (Agamaite, 1997), highlighting a few important trends (Fig. 13). First, most temporal and amplitude parameters for the four major call types show SD values near or below 0.5, while spectral parameters often show higher SD values (>0.5). Secondly, the temporal and amplitude parameters of the Phee show a higher SD for adults than for infants, as do most of the spectral parameters. The standard deviations for spectral parameters of the Trill and Trillphee cluster closely, with infant SD being higher in Trillphee parameters and adult SD being higher for most parameters of the Trill. Spectral parameters for the Twitter show a highly variable SD, with the majority being higher for the infants than for adults.

4. Individual differences

We performed PCA for all measured parameters of each call type, comparing infant vocalizations (5 weeks of age) with late juvenile vocalizations (25 weeks of age) from the same group of animals (Fig. 14). For all call types, the first principal component (PC1) consists largely of frequency pa-

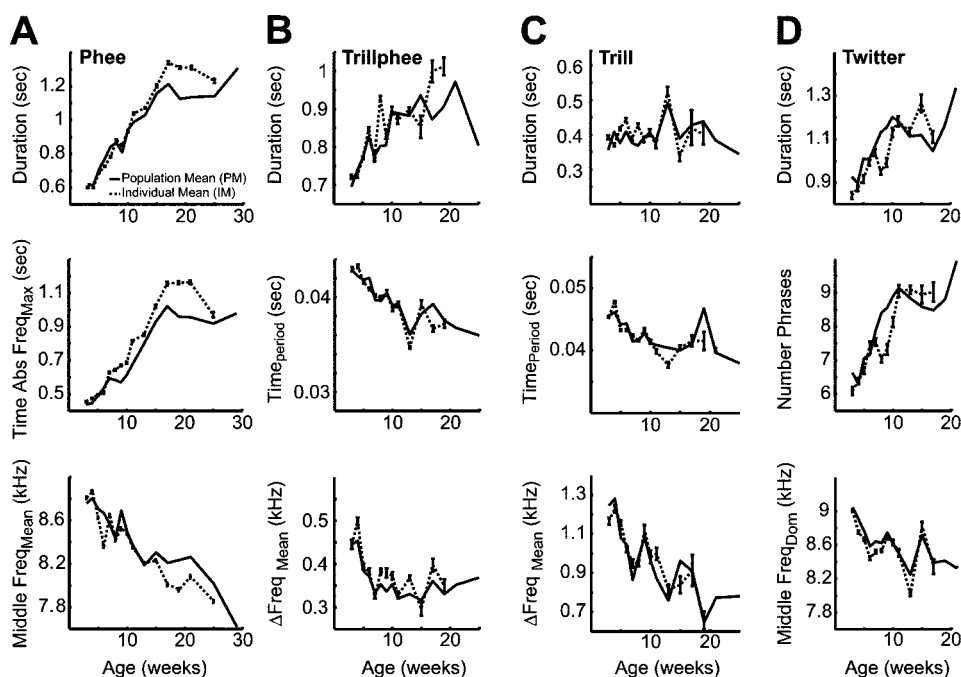


FIG. 10. Mean data from discrete age time points are plotted for select parameters of the four major call types: (A) Phee; (B) Trillphee; (C) Trill; (D) Twitter. Parameter value is on the y axis and age (in weeks) is on the x axis. The solid black line is the PM, and the dotted black line is the IM with standard error.

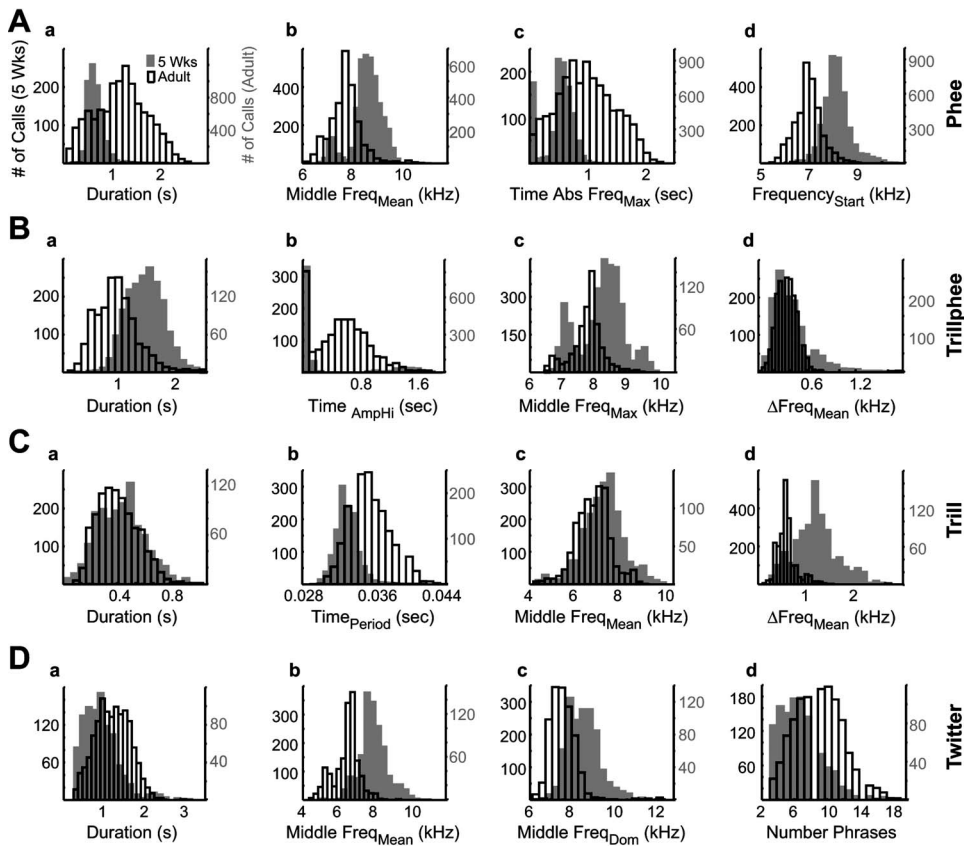


FIG. 11. Comparison between vocalizations of infant (5 weeks) and adult marmosets. Representative parameters of the four major call types: (A) Phee; (B) Trillphee; (C) Trill; (D) Twitter, from all vocalization samples are shown as histograms. Filled grey bars are the 5-week data and open bars are the adult data. The parameter value is on the x axis, the number of calls for 5-week-old infants is on the left y axis, and the number of calls for the adults is on the right y axis.

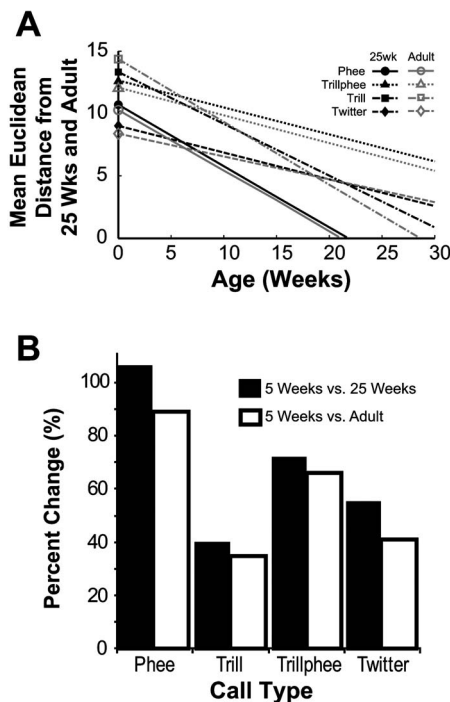


FIG. 12. Measures of parameter stabilization of animals in this study. (A) Mean Euclidean distance [as defined in Eq. (1)] (based on 18 frequency parameters) between infant marmosets at different ages and themselves at 25 weeks (solid lines and black symbols), and between infant marmosets and other adult animals (dashed line and open symbols), respectively, are plotted for the four major call types. (B) Percent change of the Euclidean distance measure between 5 weeks and 25 weeks (filled bar) or 5 weeks and adults (open bar).

parameters (black bars) and temporal parameters (grey bars) became more significant in the second principal component (PC2). In the Phee, PC1 accounts for nearly 70% of the variance. PC1 accounts for nearly 50% of the variance in the Trillphee, and, for the Trill, it accounts for nearly 40%. For the Twitter, the most acoustically complex of the four major call types, the PCA shows that PC1 and PC2 account for less variance than their counterparts for other three call types. Both PC1 ($\sim 28\%$ of variance) and PC2 ($\sim 22\%$ of variance) were dominated by frequency parameters, though to a lesser degree than those of other call types.

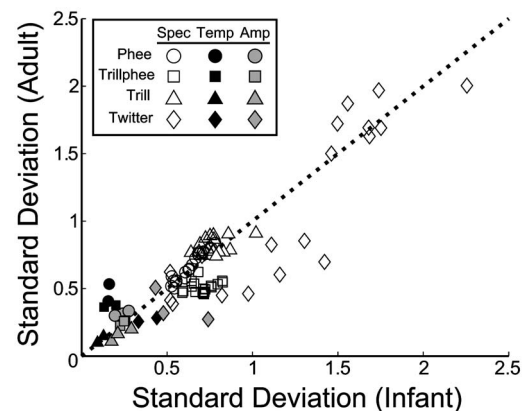


FIG. 13. Standard deviations of all measured parameters are plotted for 5-week-old infant marmosets (x axis) vs adults (y axis) for the four major call types. The diagonal line has a slope of 1. Symbol shape indicates the call type (circle: Phee; square: Trillphee; triangle: Trill; diamond: Twitter) and shading indicates the class of parameter (open: spectral; black: temporal; grey: amplitude).

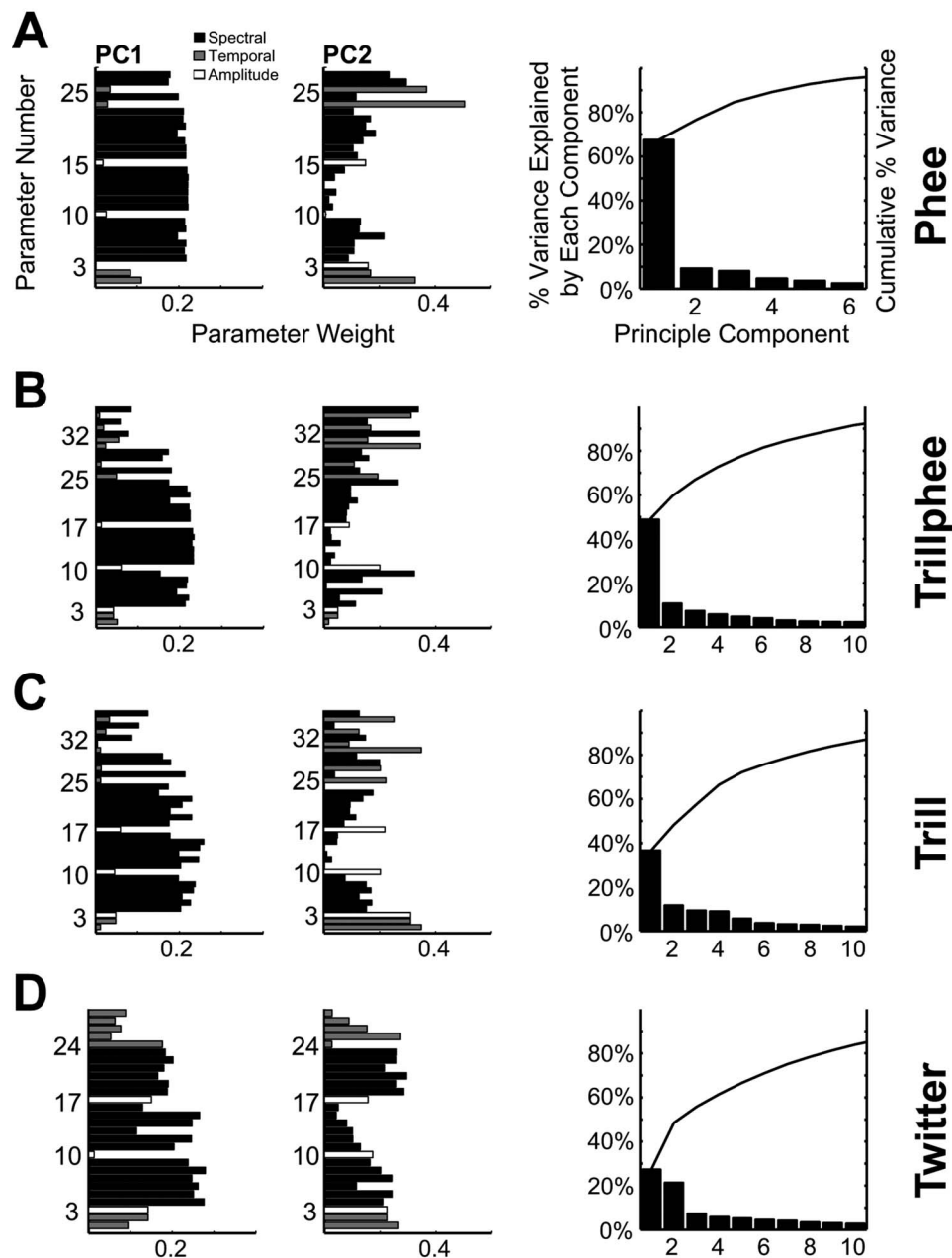


FIG. 14. Principal component analysis comparing vocalization parameters of infant (3–6 weeks) with late juvenile (19–25 weeks) for the four major call types: (A) Phee; (B) Trillphee; (C) Trill; (D) Twitter. The weights of the first (left) and second (center) principal components (PC1, PC2) are plotted in x axis number is shown on the y axis. Black bars indicate spectral parameters, grey bars indicate temporal parameters, and open bars indicate amplitude parameters. The percent of variance explained by each principal component, for principal components 1–10, is plotted on the right, with the line indicating the cumulative percent variance.

A PCA-based Discriminability Index was computed for each call type using “young” (all calls from age 3–6 weeks) versus “old” (all calls from 19 to 25 weeks) and is plotted in Fig. 15. The percent correct discrimination attributed to chance falls at 20% or 25% (dotted line on each plot), depending on the number of animals included. While all call types show above-chance discriminability between individuals at both designated age clusters, there is little change in the discrimination index for the Trillphee [Fig. 15(b), $p=0.978$] and Trill [Fig. 15(c) $p=0.872$], while it decreases significantly for the Phee [Fig. 15(a), $p<0.05$] and increases only slightly for the Twitter [Fig. 15(d), $p=0.435$] when animals became older.

5. Twin-pair comparison

The average slopes of parameter change over time for each set of twins were compared. An analysis of covariance (ANCOVA) for all parameters for each call type was performed and twins were considered significantly different for a particular parameter if $p<0.05$. In Table II, we show only significant parameters and an 18-parameter frequency analysis, which takes into account all 18 spectral parameters common across the four major call types. For the Phee call, parameters showing significant difference included duration, the time of maximum amplitude, the ending frequency, and the 18-parameter frequency distance measure. For the Trill-

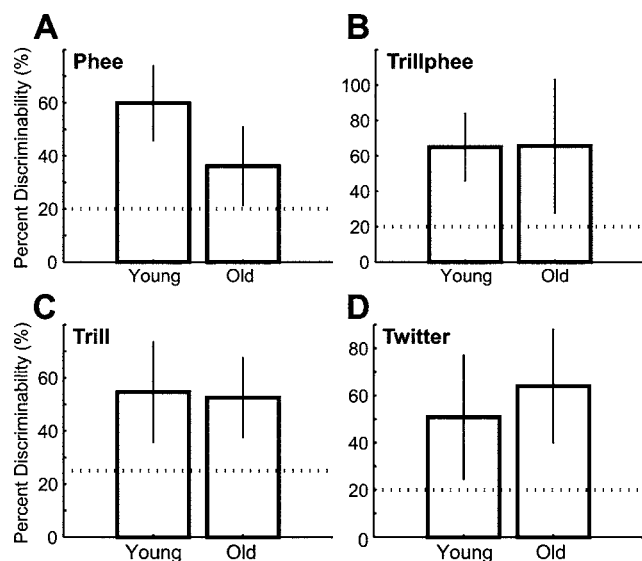


FIG. 15. Discriminability analysis based on the first two principal components (PC1 and PC2) to discriminate between individual marmosets in “young” (3–6 weeks) or “old” (19–25 weeks) groups, respectively, for the four major call types: (A) Phee; (B) Trillphee; (C) Trill; (D) Twitter. The dotted line on each plot indicates the percent correct attributed to chance (equal to one over the number of animals used in each analysis).

phee, only duration was significantly different. For the Trill, the starting frequency and frequency modulation depth differed significantly among twin pairs, and no significant difference was found in selected parameters for the Twitter.

6. Gender comparison

All collected vocalizations of the four major call types from male and female subjects were grouped accordingly and compared at two time points, at 5-weeks and at an old age. For the Phee and Trillphee, we were able to perform the analysis at 25 weeks and 21 weeks, respectively. Due to drastically decreased vocal output by females at older ages, gender differences could only be analyzed with sufficient sample size up to 13 weeks for the Trill and 19 weeks for the Twitter. Figure 16 shows some example results of this analysis. Example Phee parameters are shown from 5 and 25 weeks, Trillphee from 5 and 21 weeks, Trill from 5 and

TABLE II. ANCOVA data sample for twin pairs. Statistically significant and 18-frequency parameters (center column) are shown for each of the four call types (left column) with results from ANCOVA tests. p -values are in the right column and numbers in bold indicate when twin pairs were found to be significantly different ($p < 0.05$).

Call type	Parameter	p
Phee	Duration	0.0004
	Time _{AmpHi}	0.00001
	Frequency _{End}	0.0078
	18-Parameter Frequency Distance	0.0001
Trillphee	Duration	0.0394
	18-Parameter Frequency Distance	0.1314
Trill	Frequency _{Start}	0.0323
	Δ Freq _{Mean}	0.0117
	18-Parameter Frequency Distance	0.0759
Twitter	18-Parameter Frequency Distance	0.0932

13 weeks, and Twitter from 5 and 19 weeks to illustrate the shift in distribution over age, as well as the presence or absence of separation by gender.

Unexpectedly, for the Trillphee [Fig. 16(b)], the mean duration increased in males, while it decreased in females. Another surprising result was the change in the mean of the middle mean frequency, which rose for females while dropping in males. For the Trill [Fig. 16(c)], the mean modulation depth decreased significantly, and the difference between means for males and females increased. Figure 16(d) shows age-related trends among males and females for the Twitter. In general, frequency parameters do not show large changes over time but do show differences between males and females.

In a gender-based ANCOVA analysis (Table III) of the magnitude and direction of change for select parameters for each call type, males and females show little difference in the Phee calls. Some significant differences found were in the absolute maximum frequency in the Trill, the start frequency in the Trill and Trillphee, and the duration and number of phrases in the Twitter. We also computed the 18 frequency-parameter distance measure between males and females for each call type (Table III). The only significant difference in this measure was found for the Twitter call.

IV. DISCUSSION

A. Summary of findings

The present study showed that infant marmosets exhibited a high rate of calling, the use of many call types in the absence of context, as well as possessing infant-specific call types and features such as the Cry and Twitter-hook. Based on our detailed acoustic analysis, the vocalizations of infant marmosets differed from those of adults in both the spectral and temporal domains. In addition, these features changed gradually through development in the direction of the feature distributions of the adult calls. One could argue that a gradual change in vocalizations is more likely to (but not necessarily) result from experience-based plasticity or learning. Had we observed no changes or an abrupt change shortly after birth, such a possibility is much lower. One would have to argue for a genetic or innate explanation in that case.

Because individual recognition based on vocalizations has been shown across many species of animal, we compared changes in our defined parameters for each call type across individuals, twin-pairs, and genders to determine which parameter(s) may be used by marmosets to identify conspecifics as being from the same family group, of a particular gender, or as individuals. Certain features in each call type developed differently depending on the gender and family of origin. When taken as a sum of differences in one animal, based on gender and family group, it is possible to see how individual characteristics could be developing as well. Additionally, the results of the ANCOVA analysis of twin pairs and gender suggest that many of the measured parameters for each call type develop similarly across all infant marmosets. While the starting and ending points of acoustic parameters we measured may be different for *individual* infants, trends

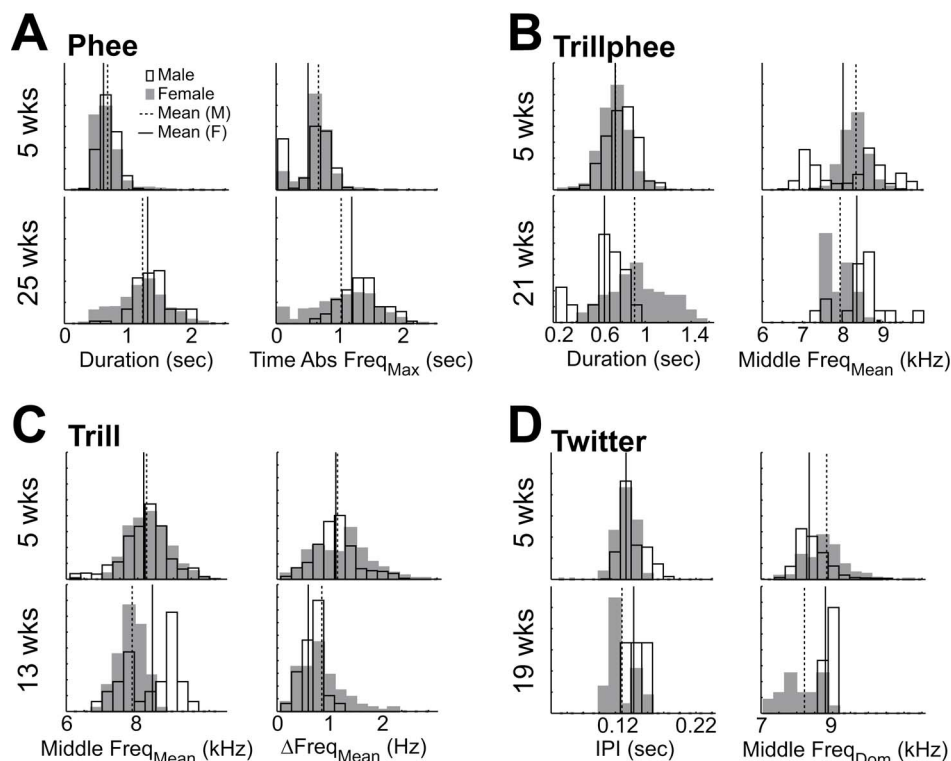


FIG. 16. Gender comparison for two example parameters from the four major call types: (A) Phee; (B) Trillphee; (C) Trill; (D) Twitter, at 5 weeks and an older age (25 weeks for Phee, 21 weeks for Trillphee, 13 weeks for Trill, and 19 weeks for Twitter) is plotted in histogram format. Parameter value is on the x axis. Filled grey bars are females and open bars are males. The mean of each distribution is plotted as a vertical solid line for females and a vertical dashed line for males.

in infant marmoset vocalization parameters follow the same approximate slope during development.

B. Limitations of the present study

Definitions of parameters for each call type were based on the acoustic characterization and statistics of adult marmoset calls. Because it is difficult to determine *a priori* which parameters are involved in the neural coding or behavior of vocalization perception and production in common marmosets, we analyzed as many as possible and determined which were contributing most to variability between genders, twin pairs, and individuals. Additionally, there may be parametric characteristics in infant calls associated with a certain behavior, but are less meaningful in adult vocalizations. This study focused on the development of infant marmoset vocalizations and quantified the progress of each infant in fitting

adult characteristics, and presumably in developing individuality. For determining the roles of infant-specific behaviors in the social dynamic groups, a different type of study is warranted.

In this study, no attempt was made to classify calls within one call type into subtypes. Previous studies of primate species have shown that there is a continuum along vocalization acoustics, making calls less stereotyped as one expands to finer detail in parameters. For example, marmosets use calls fitting the characteristics of a Phee in different contexts and vary the parameters depending on behavioral context (e.g., multipart Phees and low-amplitude Phees). For the purpose of the present study, which emphasized acoustic analysis as the first step towards a comprehensive description of vocal development in this species, we did not analyze behavioral context. This allowed for a large sample size of vocalizations to be collected and analyzed in each of the four major call types studied, and their statistics to be obtained. There are likely to be subtle acoustic differences within call types depending on the behavioral state of the animal during vocal production. Future studies incorporating proper behavioral analyses would be required to clarify these differences.

Although the recording setup used in the present study was vital to obtain high quality samples for detailed acoustic analyses, the physical isolation of infants during recording sessions constitutes an interruption of their normal social environment, even though they maintained visual, auditory, and olfactory contact with their family group. We suspect that such interruptions may have affected the contextual informa-

TABLE III. ANCOVA data sample for gender. Same format as Table II.

Call type	Parameter	p
Phee	18-Frequency Parameter Distance	0.3752
Trillphee	Frequency _{Start}	0.0316
	18-Frequency Parameter Distance	0.6547
Trill	Absolute Freq _{Max}	0.0051
	Frequency _{Start}	0.003
	18-Frequency Parameter Distance	0.5047
Twitter	Duration	0.0026
	Number Phrases	0.0023
	18-Frequency Parameter Distance	0.0438

tion of produced calls more so than their acoustic structures, which were the focus of the analyses in this paper. Presumably, vocal behavior of a normal adult marmoset is different in isolation than it is in a social setting such as group housing, and vocal output would skew towards calls typically used in that context. This is apparent in the older animals in this study, when there is a substantial drop in the overall vocal output and the Phee dominates the recorded samples. This caveat is also an advantage; it is very telling about contextual awareness in infant marmosets. Whether spontaneous vocalization in infants is for the purpose of practice, as seen in birds or human infants, remains to be proven, but there is a clear time point when infant marmosets decrease the number of spontaneous vocalizations and it coincides with an increasing number of vocal exchanges with conspecifics.

C. Can we establish a non-human primate model of vocal production learning?

The nonhuman primate is, perhaps, the most desirable model for human speech and language development. It would be an incredible advancement of the field if it could be proven that they display similar capacities, because of the extremely close ties both genetically and anatomically. While the specific question of vocal learning was not addressed in this study, our results will reopen the debate of whether non-human primates exhibit ontogenetic development in the production of their vocalizations.

Recently, Egnor and Hauser (2004) addressed an apparent paradox in recent studies showing auditory feedback-dependent plasticity of adult primate vocalizations, despite older studies showing no plasticity in the development of vocal production in juvenile nonhuman primates, and they appropriately called for a more careful investigation of these contradictory findings. This study aims to lay groundwork for future studies into the vocal development of common marmosets, as well as other nonhuman primate species. We show that with strong acoustic recording and analysis, there are discrete changes in the vocalizations produced by developing marmosets.

An acoustic analysis such as the one presented here is only the first step in a study of whether vocal learning occurs, satisfying the primary requirement that significant ontogenetic changes do, in fact, occur. Determining what changes are due to maturation of the vocal tract and general growth requires experiments that perturb normal development. Behavioral and neural studies should be conducted in conjunction with detailed acoustic analysis techniques, such as those presented here, in animals with abnormal auditory feedback, such as those deafened early in life or congenitally, ear plugged during developmental stages, noise-raised, or with implanted earphone filters which could filter certain frequencies or shift feedback in pitch and/or time. Other methods of perturbation could include normally-hearing infants born to mute or vocally abnormal parents and raised isolated from normal conspecifics, infants with an altered vocal apparatus, or infants cross-fostered with a compatible species that has very different vocalizations. Behavioral studies, acoustic analysis, and other technology such as electrophysi-

ology, functional imaging, and immediate early gene induction could contribute to answering the final question of whether there is any auditory feedback-dependent plasticity in vocal production development of nonhuman primates.

ACKNOWLEDGMENTS

We gratefully acknowledge Dr. Cory Miller for his comments and suggestions on this manuscript. This work was supported by NIH Grant No. DC005808 (X.W.).

- Agamaite, J. A. (1997). "A quantitative characterization of the vocal repertoire of the common marmoset," Master's thesis, Johns Hopkins University, Baltimore, MD.
- Biben, M. and Bernhards, D. (1995). "Vocal ontogeny of the Squirrel Monkey, *Saimiri boliviensis peruviansis*," in *Current Topics in Primate Vocal Communication*, edited by Zimmerman, E., Newman, J. D., and Jurgens, U., (Plenum, New York), pp. 99–120.
- Egnor, S. E. R. and Hauser, M. D. (2004). "A paradox in the evolution of primate vocal learning," *Trends Neurosci.* **27**, 649–654.
- Epple, G. (1968). "Comparative studies on vocalization in marmoset monkeys (*Hapalidae*)," *Folia Primatol.* **8**, 1–40.
- Fischer, J., Cheney, D. L., and Seyfarth, R. M. (2000). "Development of infant baboons' responses to graded bark variants," *Proc. R. Soc. London, Ser. B* **267**, 2317–2321.
- Fischer, J. (2003). "Developmental modification in the vocal behavior of non-human primates," in *Primate Audition: Ethology and Neurobiology*, edited by A. A. Ghazanfar (CRC Press, Boca Raton), pp. 109–125.
- Hammerschmidt, K., Newman, J. D., Champoux, M., and Suomi, S. J. (2000). "Changes in rhesus macaque 'coo' vocalizations during early development," *Ethology* **106**, 873–886.
- Hammerschmidt, K., Freudenstein, T., and Jürgens, U. (2001). "Vocal development in squirrel monkeys," *Behaviour* **138**, 1179–1204.
- Inoue, M. (1988). "Age gradations in vocalization and body weight in Japanese monkeys (*Macaca fuscata*)," *Folia Primatol.* **51**, 76–86.
- Kalin, N. H. and Shelton, S. E. (1998). "Ontogeny and stability of separation and threat-induced defensive behaviors in rhesus monkeys during the first year of life," *Am. J. Primatol.* **44**, 125–135.
- Masataka, N. (1993). "Effects of contingent and noncontingent maternal stimulation on the vocal behaviour of three- to four-month-old Japanese infants," *J. Child Lang.* **20**, 303–312.
- McCowan, B., Franceschini, N. V., and Vicino, G. A. (2001). "Age differences and developmental trends in alarm peep responses by squirrel monkeys (*Saimiri sciureus*)," *Am. J. Primatol.* **53**, 19–31.
- Merker, B. and Cox, C. (1999). "Development of the female great call in *Hylobates gabriellae*: A case study," *Folia Primatol.* **70**, 97–106.
- Miller, C. T. and Wang, X. (2006). "Sensory-motor interactions modulate a primate vocal behavior: Antiphonal calling in common marmosets," *J. Comp. Physiol., A* **192**, 27–38.
- Moody, M. I. and Menzel, E. W., Jr. (1976). "Vocalizations and their behavioral contexts in the tamarin *Saguinus fuscicollis*," *Folia Primatol.* **25**, 73–94.
- Newman, J. D. and Symmes, D. (1974). "Vocal pathology in socially deprived monkeys," *Dev. Psychobiol.* **7**, 351–358.
- Newman, J. D. (1995). "Vocal Ontogeny in Macaques and Marmosets: Convergent and Divergent Lines of Development," in *Current Topics in Primate Vocal Communication*, edited by E. Zimmerman, J. D. Newman, and U. Jurgens (Plenum, New York), pp. 73–97.
- Norcross, J. L. and Newman, J. D. (1993). "Context and gender-specific differences in the acoustic structure of common Marmoset (*Callithrix jacchus*) Phee calls," *Am. J. Primatol.* **30**, 37–54.
- Norcross, J. L. and Newman, J. D. (1997). "Social context affects phee call production by nonreproductive common marmosets (*Callithrix jacchus*)," *Am. J. Primatol.* **43**, 135–145.
- Owren, M. J., Dieter, J. A., Seyfarth, R. M., and Cheney, D. L. (1992). "Food" calls produced by adult female rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques, their normally-raised offspring, and offspring cross-fostered between species," *Behaviour* **120**, 218–231.
- Owren, M. J., Dieter, J. A., Seyfarth, R. M., and Cheney, D. L. (1993). "Vocalizations of rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques cross-fostered between species show evidence of only limited modification," *Dev. Psychobiol.* **26**, 389–406.

- Schrader, L. and Todt, D. (1993). "Contact call parameters covary with social context in common marmosets, *Callithrix j. jacchus*," *Anim. Behav.* **46**, 1026–1028.
- Seyfarth, R. M. and Cheney, D. L. (1986). "Vocal development in vervet monkeys," *Anim. Behav.* **34**, 1640–1658.
- Seyfarth, R. M. and Cheney, D. L. (1997). "Some general features of vocal development in nonhuman primates," in *Social Influences on Vocal Development*, edited by C. T. Snowdon and M. Hausberger (Cambridge University Press, U.K.), pp. 249–273.
- Snowdon, C. T. (1997). "Affiliative processes and vocal development," *Ann. N.Y. Acad. Sci.* **807**, 340–351.
- Snowdon, C. T., Elowson, A. M., and Roush, R. S. (1997). "Social influences on vocal development in New World Primates," in *Social Influences on Vocal Development*, edited by C. T. Snowdon and M. Hausberger (Cambridge University Press, U.K.), pp. 234–248.
- Snowdon, C. T. and Elowson, A. M. (2001). "'Babbling' in pygmy marmosets: Development after infancy," *Behaviour* **138**, 1235–1248.
- Stevenson, M. F. and Poole, T. B. (1976). "An ethogram of the common Marmoset (*Callithrix jacchus jacchus*): General behavioural repertoire," *Anim. Behav.* **24**, 428–451.
- Winter, P. (1969). "The variability of peep and twit calls in captive squirrel monkeys (*Saimiri sciureus*)," *Folia Primatol.* **10**, 204–215.
- Winter, P., Handley, P., Ploog, D., and Schott, D. (1973). "Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation," *Behaviour* **47**, 230–239.