



The covariation of acoustic features of infant cries and autonomic state



Adam Michael Stewart^a, Gregory F. Lewis^a, Keri J. Heilman^a, Maria I. Davila^a, Danielle D. Coleman^a, Stephanie A. Aylward^a, Stephen W. Porges^{a,b,*}

^a Brain-Body Center, Department of Psychiatry, University of Illinois at Chicago, 1601 W. Taylor Ave., Chicago, IL 60612, USA

^b Department of Psychiatry, University of North Carolina, Chapel Hill, NC 27599, USA

HIGHLIGHTS

- Vocal prosody and the heart are regulated by a shared neural pathway.
- A novel automated tool was applied to objectively characterize vocal prosody.
- Vocal prosody is a sensitive index of autonomic activity in human infants.
- Specific acoustic features conveying vocal prosody covary with physiological state.
- Faster heart rate was associated with restricted modulation (i.e., less prosody) of acoustic features.

ARTICLE INFO

Article history:

Received 26 December 2012

Received in revised form 19 June 2013

Accepted 22 July 2013

Keywords:

Infant vocalizations

Prosody

Polyvagal Theory

Autonomic nervous system

Heart rate

Respiratory sinus arrhythmia

ABSTRACT

The evolution of the autonomic nervous system provides an organizing principle to interpret the adaptive significance of physiological systems in promoting social behavior and responding to social challenges. This phylogenetic shift in neural regulation of the autonomic nervous system in mammals has produced a neuroanatomically integrated social engagement system, including neural mechanisms that regulate both cardiac vagal tone and muscles involved in vocalization. Mammalian vocalizations are part of a conspecific social communication system, with several mammalian species modulating acoustic features of vocalizations to signal affective state. Prosody, defined by variations in rhythm and pitch, is a feature of mammalian vocalizations that communicate emotion and affective state. While the covariation between physiological state and the acoustic frequencies of vocalizations is neurophysiologically based, few studies have investigated the covariation between vocal prosody and autonomic state. In response to this paucity of scientific evidence, the current study explored the utility of vocal prosody as a sensitive index of autonomic activity in human infants during the Still Face challenge. Overall, significant correlations were observed between several acoustic features of the infant vocalizations and autonomic state, demonstrating an association between shorter heart period and reductions in heart period and respiratory sinus arrhythmia following the challenge with the dampening of the modulation of acoustic features (fundamental frequency, variance, 50% bandwidth, and duration) that are perceived as prosody.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

The phylogenetic shift in neural regulation of the autonomic nervous system (ANS) in vertebrates has produced a myelinated efferent vagal pathway unique to mammals, which can rapidly regulate visceral state and facilitate social engagement [1–3]. The brainstem origin of this motor-specific vagal pathway shares evolutionary origins with the special visceral efferent pathways traveling through five cranial nerves (trigeminal, facial, glossopharyngeal,

vagus, and accessory), which innervate somatic muscles typically not associated with the neurophysiology of the ANS. In addition to its innervation of smooth and cardiac muscles, the mammalian vagus is neuroanatomically linked to the striated muscles of the face and head that regulate social engagement, intimately involved with behaviors such as facial expression, hearing, and vocalization, that contribute to the complex repertoire of both social expressions (i.e., signaling) and social experiences (i.e., receiving) [1–4].

Mammalian vocalizations are part of a social communication system that evolved to communicate and to signal a danger or threat to conspecifics, with many mammalian species modulating their acoustic features of vocalizations to signal states of fear, aggression, pain, and hunger [5–7]. Human infant crying, in particular, serve to elicit parental caregiving, with variations in their cry features known to

* Corresponding author at: Department of Psychiatry, University of North Carolina School of Medicine, 387 Medical School Wing D, Campus Box 7160, Chapel Hill, NC 27599-7260, USA. Tel.: +1 919 843 2220.

E-mail address: stephen_porges@med.unc.edu (S.W. Porges).

affect adults' perceptions and physiologic responses [8–11]. At the clinical level the acoustic features of infants have been shown to exhibit a wide variance relative to the intensity or degree of a particular challenge and/or medical status [8,12,13]. For example, previous reports have shown a relationship between acoustic characteristics of infant cries and diagnoses related to neurological damage, prematurity, medical conditions, pain (e.g., circumcision), sudden infant death syndrome, and substance exposure during pregnancy [8,14–19]. In particular, the frequency of infant cries may be interpreted as indicators of health status and [20–23], and can indicate distress when infants exhibit high frequency shrill cries [24–27].

Importantly, this covariation between physiological state and the acoustic frequencies of vocalizations is neurophysiologically based on a shared neural pathway, as suggested by the Polyvagal Theory [5]. Prior reports have demonstrated the relationship between variations in infant cry prosody and infant neurobehavioral organization and autonomic activity [28]. For example, higher cry threshold and shorter cry duration have been shown to be predictive of disrupted autonomic regulation [28]. However, the parallel between the acoustic features characterized as vocal prosody (i.e., the cluster of vocal features such as pitch modulation and frequency change) has not been quantified in detail, nor has its parallel with autonomic function been systematically described. To date, only the frequency of vocalization has been correlated to cardiac vagal tone (i.e., respiratory sinus arrhythmia) and then only in healthy, although severely stressed, infants [29]. Here we expand these previous observations to demonstrate the covariation of several prosodic features with autonomic state in human infants.

1.1. Role of vocalizations

The physiological correlates of vocal prosody have been understudied, although in clinical settings vocal prosody has been assumed to be related to health. Most investigations have focused on vocalization as a reflection of neurobehavioral dysregulation, often in relation to a specific disorder [30–32], contextual features (e.g., danger) [33,34], or transitory psychological state (e.g., pain) [35,36] and with little scientific interest in a potential relation between prosody and autonomic state. Consistent with the Polyvagal Theory [5], the current study assumes that the frequency modulations within vocalizations are an effective and efficient form of vocal communication with an adaptive function, capable of conveying the physiological state of the signaler to other members of its species. This 'automatic' communication of the signaler's state is hypothesized to facilitate reproductive, parenting, and social behaviors by indicating when the organism is safe to approach [5].

Prosodic vocalization characteristics are generally expanded during positive social interactions among humans, and reduced during stress or illness [29,37,38]. For example, distressed infants often demonstrate a high-pitched cry with little frequency modulation and shortened duration [16,29]. Consistent with the polyvagal hypothesis, the decreased neural tone seen in distressed or unhealthy infants is believed to reduce the inhibitory effect on the heart and bronchi and the contraction of laryngeal muscles, thereby producing dramatic increases in both heart rate and respiration rate as well as the fundamental frequency of the cry [5,29,39,40]. This convergence of cardiac vagal tone and vocal prosody has also been demonstrated in healthy, though severely stressed, newborns following circumcision. Cardiac vagal tone, measured by respiratory sinus arrhythmia (RSA), was significantly reduced, and individual differences in resting cardiac vagal tone were shown to correlate with the pitch of the infants' high frequency pain cries [29].

1.2. Shared neurophysiology underlying vocal prosody and autonomic regulation

The evolutionary development of the striated muscles necessary for mammalian vocalizations paralleled the evolutionary changes in vagal

regulation of the heart [41]. Ultimately, this phylogenetic convergence of several neural circuits has served to form an integrated functional social engagement system [4,42,43]. Importantly, the same brainstem structures are involved in both the regulation of the heart rate via the myelinated branch of the vagus (i.e., cardiac vagal tone measured via respiratory sinus arrhythmia) and vocalizations via the laryngeal and pharyngeal muscles [44,45]. Thus, the vagal output to the laryngeal and pharyngeal muscles reflected in the prosodic features of vocalization may mirror the vagal influence to the heart. Therefore, procedures that disrupt homeostatic processes may result in both depressed cardiac vagal tone and vocalizations characterized by a higher pitch and less prosody [5]. Moreover, a chronically depressed cardiac vagal tone would reflect, on an individual level, poor homeostasis and a neurophysiological vulnerability to a challenge.

The brain substrates mediating vocalizations remain relatively conserved across most mammalian species [46], and with the exception of some cortical influences, the brain structures reside primarily in the subcortical forebrain — namely limbic structures and the hypothalamus, as well as the periaqueductal gray of the midbrain, leading to the motor neurons that innervate the larynx [46,47]. Regulation of special visceral efferents to the laryngeal and pharyngeal muscles involved in vocalization, as well as general visceral efferents producing a rhythmic oscillation in heart associated with breathing (i.e., RSA), stem from the myelinated vagal pathway originating in the nucleus ambiguus (NA).

Vocalizations represent direct central regulation of laryngeal and pharyngeal muscles, with the neural regulation of these muscles occurring in the same brainstem nuclei involved in regulation of myelinated vagal pathways to the heart [46,47]. In particular, specific vocalization features are plausibly linked to the neural regulation of the laryngeal and pharyngeal muscles. For example, a positive relationship between the height of the larynx muscle and vocal fundamental frequency exists, with vocal pitch changing relative to changing larynx height (i.e., the relaxation and contraction of the muscle) [48,49]. Thus, rapid change in fundamental frequency during a vocalization represents modulation of neural tone to the laryngeal muscles. The laryngeal muscles provide active opening of the glottis (i.e., the combination of the vocal cords and the space in between the folds) to enhance the flow during inspiration, and partial closing to reduce air flow and increase subglottal pressure during expiration. During expiration, the posterior cricoarytenoid (PCA) muscle (regulated by the recurrent laryngeal branch of the vagus) is phasically active while activity of the cricothyroid (CT) muscle (regulated by superior laryngeal nerve) tends to increase. Likewise, the thyroarytenoid (TA) muscle (regulated by superior laryngeal nerve) is also more active during expiration. The TA is comprised of two divisions — with the external division (TA-X) adducting the vocal fold, and the vocalis division (TA-V) modulating sound quality [50].

The activity of the laryngeal and pharyngeal muscles parallel the alternating increase and decrease in myelinated vagal output to the heart during exhalation and inhalation, respectively (Table 1). This occurs because respiration modulates vocalizations and “gates” the efferent action of the myelinated vagus on the sino-atrial node [51]. Moreover, with the laryngeal and pharyngeal muscles and heart both

Table 1

Laryngeal regulation of the glottis parallels the alternating increase and decrease in myelinated vagal input to the heart during exhalation and inhalation (i.e., RSA).

Inhalation	Exhalation
Less vagal output	More vagal output
Laryngeal muscles abduct	Laryngeal muscles adduct
Posterior cricoarytenoid muscle (via recurrent laryngeal nerve)	Thyroarytenoid muscle (via recurrent laryngeal nerve)
Cricothyroid muscle (via superior laryngeal nerve)	
Glottis opens	Glottis closes

innervated by a common pathway originating in the NA, and subject to the same afferent regulation via the nucleus tractus solitarius (NTS), a shared modulation is not surprising [52,53]. The common features between the brainstem areas regulating the heart and the structures producing vocalization provide the neurophysiological justification to investigate the covariation between acoustic features of vocalizations and vagal influences to the heart.

Given the shared neural regulation between vocalization and autonomic state, this study investigates the covariation of several features of vocal prosody with autonomic measures dependent on vagal regulation of the heart. In the current study, a recently developed automated tool for objectively characterizing the features of vocalizations was applied. Vocalization features were then assessed and correlated with RSA and heart rate (i.e., heart period).

2. Methods

2.1. Subjects

Vocalization analysis was performed on audio files ($n = 75$) previously recorded in a study investigating the autonomic response profile of infants at 6-months to the Still Face procedure (see below). The audio files represent a subset of 75 infant participants from the above study who vocalized during the procedures. The infants were recruited between 6 and 10 weeks. At 6 months they received a developmental assessment and participated in a laboratory experiment to evaluate autonomic reactivity.

2.2. Still Face paradigm

A modified Face-to-Face Still-Face (FFSF) paradigm [54] was used for this study. While also providing information regarding the physiological reactivity to a challenge and its contingency on the ability to regulate visceral state, the use of this paradigm in the context of this experiment primarily served as a means to evoke vocalizations through a social challenge. In the modified protocol, the infant was placed in a high chair, and the experimenter was seated so that her face was at the infant's eye level and approximately 50 cm from the infant. The mother was absent from the research room during the protocol. The social challenge session consisted of three 2-min conditions: Free Play 1, Still Face and Recovery (Free Play 2) trials. During Free Play 1 the experimenter interacted with an infant in a playful manner. During Still Face the experimenter stopped smiling and expressed a blank face while looking at the infant. The experimenter did not move her head, or use touch or speech to engage the infant. The Still Face trial was terminated earlier than 2 min if the infant became distressed. During Recovery the experimenter resumed smiling and used speech and gestures to re-engage the infant. Physiological monitoring with the LifeShirt (see next section) was conducted continuously through the social challenge to extract autonomic measures during each of the 2-min segments.

2.3. Collection of autonomic measures

The LifeShirt System (VivoMetrics, San Diego, CA), a non-invasive, continuous ambulatory monitoring system, was used to collect beat-to-beat heart periods (i.e., the time between successive heart beats). The LifeShirt consists of a comfortable garment with an array of embedded sensors, a handheld computer, and a sophisticated software package. Measures of respiratory frequency were quantified for each phase. Three Ag/AgCl ECG self-adhering electrodes (Covidien, Mansfield, MA) were placed on the participant by a female laboratory assistant and connected to the LifeShirt System.

2.4. Quantification of autonomic variables

Heart period data, derived from the ECG signal, were then edited off-line with CardioEdit software (Brain-Body Center, Chicago, IL). RSA was defined as the variance in the beat-to-beat heart rate pattern in the frequency band associated with spontaneous breathing. In the current study RSA was calculated using CardioBatch software (Brain-Body Center, Chicago, IL) and operationally defined as the variance in the frequency band between .24 and 1.04 Hz using the Porges–Bohrer algorithms incorporated in CardioBatch [55–57]. RSA and heart period (i.e., the time in ms between successive heart beats) were calculated in sequential 30 s epochs. Changes in RSA and heart period were also calculated by subtracting values derived during Free Play 1 from values derived during the Still Face challenge (i.e., SF-FP1). Averages of the 30 s epochs within each condition (i.e., Free Play, Still Face) and the changes in heart period and RSA between the conditions were used in the analyses.

2.5. Quantification of acoustic measures

Audio files were extracted from the video recordings of the modified FFSF protocol. Individual vocalizations were manually extracted with Adobe Audition (Adobe Systems, San Jose, CA) and then uploaded into a LabView-based program for quantification of the prosodic features. Only audio files demonstrating one or more vocalization bouts during the 2 min SF challenge were analyzed. The vocalizations ranged from stereotypical “coos” to robust bouts of crying and were defined as periods of audible sound during expiration that lasted for at least 100 ms. Sounds primarily due to biological functions (e.g., gurgling) were excluded. Bouts of crying were counted as individual vocalizations when a change from expiration to inspiration occurred.

Several variables were operationally defined to index features assumed to contribute to the prosodic features of mammalian vocalizations. Prosody, representing the modulation of vocalization frequencies, was deconstructed into several quantifiable variables. First, the fundamental frequency of the vocalization was quantified. The fundamental is, when viewed in a spectrograph, the first energy peak. The fundamental frequency represents the perceptual attribute of vocal pitch. Second, the variance of the fundamental quantifies how much the fundamental pitch varies during the vocalization. Third, The 50% bandwidth reflects the frequency band in which fundamental varies between 50% and 100% of its peak energy. When viewed from a graphic perspective of an FFT, the 50% bandwidth is the horizontal distance (i.e., band of frequencies) across the first energy peak (i.e., the fundamental) at 50% of the peak maximum (visually, the horizontal measurement of width is taken at half the peak's maximum energy). Together, variance and 50% bandwidth are most closely related to pitch variability. Fourth, vocalization duration was calculated as the time in which the energy (amplitude) of the fundamental crossed a defined threshold demarcating vocalization energy from background noise.

The acoustic features were quantified using a dynamic, computation-based tool capable of rapidly detecting and characterizing the features of vocalizations developed in our laboratory. Individual vocalizations were manually extracted using Adobe Audition, and exported as individual .wav files into a LabView-based user interface. With the initiation of the LabView program, a Fourier transform was performed on each vocalization to convert the time-domain audio waveform into the frequency-domain. A peak detection process, conducted in the frequency domain, isolated the fundamental (first energy peak) and harmonics (subsequent energy peaks) of the vocalization, and quantified the associated frequency and magnitude values. Variance and duration were obtained by converting the spectrogram into a two-dimension array of coordinates representing the frequency and time domains. Thus, both measures were not based on an average or peak value of the fundamental, but rather on the fundamental frequency of the entire expiratory sound as long as it was expressed

Table 2

Mean values for acoustic and autonomic measures across all subjects ($n = 75$), as well as for male ($n = 40$) and female ($n = 35$) subjects separately (presented as mean \pm standard deviation). RSA and heart period changes based on difference between measures obtained from Free Play 1 to the Still Face procedure (i.e., SF–FP1).

	All subjects	Males	Females
RSA during FP1	3.25 \pm 0.74	3.28 \pm 0.69	3.20 \pm 0.81
HP during FP1	441.78 \pm 33.97	440.54 \pm 32.17	443.17 \pm 36.31
RSA during SF	3.14 \pm 0.91	3.08 \pm 0.73	3.21 \pm 1.07
HP during SF	409.96 \pm 39.54	406.00 \pm 35.72	414.36 \pm 43.08
RSA change	0.08 \pm 0.76	−0.20 \pm 0.75	0.06 \pm 0.76
HP change	−30.26 \pm 38.17	−32.62 \pm 37.54	−27.68 \pm 38.68
Fund. med.	794.40 \pm 294.03	821.69 \pm 334.77	763.21 \pm 233.33
Fund. min.	484.73 \pm 181.85	479.9855 \pm 168.23	490.15 \pm 196.52
Fund. max.	1471.74 \pm 858.42	1481.1475 \pm 1021.29	1460.99 \pm 613.57
BW med.	4.77 \pm 0.37	4.796 \pm 0.38	4.74 \pm 0.35
BW min.	4.20 \pm 0.38	4.21 \pm 0.43	4.17 \pm 0.32
BW max.	5.58 \pm 0.56	5.62 \pm 0.55	5.54 \pm 0.58
Var. med.	6.93 \pm 1.59	6.64 \pm 1.54	7.26 \pm 1.60
Var. min.*	5.19 \pm 1.30	4.86 \pm 1.06	5.55 \pm 1.43
Var. max.	8.49 \pm 2.10	8.21 \pm 2.18	8.81 \pm 1.96
Dur. med.	0.29 \pm 0.31	0.24 \pm 0.27	0.34 \pm 0.33
Dur. min.	0.11 \pm 0.29	0.09 \pm 0.20	0.13 \pm 0.34
Dur. max.*	1.02 \pm 0.70	0.81 \pm 0.65	1.21 \pm 0.70

"FP" = Free Play; "SF" = Still Face challenge; "RSA" = respiratory sinus arrhythmia (ms^2); "HP" = heart period (ms); "Fund." = fundamental frequency (Hz); "BW" = 50% bandwidth (Hz); "Var." = variance (Hz); "Dur." = duration (s).

* Significant sex differences (Mann–Whitney U).

over a given magnitude threshold (i.e., a threshold set above background noise). An output of all quantified acoustic variables was generated in both the user-interface as well as a .csv file for further processing in Excel.

Following quantification with the automated tool, the raw values for each acoustic variable were processed to calculate the median, minimum, and maximum values per subject. Median was used as a measure of central tendency to minimize influence of atypical distributions and outliers. Minimum and maximum values were quantified to evaluate whether certain indices may be more sensitive to changes at either the lower or higher ends of a variable's range (e.g., the high fundamental frequency demonstrated by severely stressed infants may represent the higher end of the possible frequency range.) The natural logarithms of the 50% bandwidth and variance of the fundamental were used here to ensure that the metrics complied with the assumptions for parametric analyses and were appropriate for linear regression modeling of the total prosodic range.

Table 3

Correlation between acoustic and autonomic variables for infant vocalizations. RSA and heart period changes based on difference between measures obtained from Free Play 1 to the Still Face procedure (i.e., SF–FP1).

	RSA during FP1	HP during FP1	RSA during SF	HP during SF	RSA change	HP change
Fund. med.	.048	.193	−.022	−.253*	−.104	−.444**
Fund. min.	−.170	.019	.047	.225	.180	.195
Fund. max.	.057	.210	−.146	−.368**	−.241*	−.517**
BW med.	.158	−.030	−.026	.138	−.201	.132
BW min.	.003	−.219	.094	.285*	.135	.454**
BW max.	.137	.104	−.157	−.334**	−.315**	−.410**
Var. med.	−.006	−.012	−.094	−.330**	−.076	−.296**
Var. min.	−.048	−.020	.192	.309*	.284*	.337*
Var. max.	.026	−.072	−.193	−.610**	−.217	−.518**
Dur. med.	−.192	−.124	.118	−.079	.311*	.051
Dur. min.	−.112	−.089	.244	.242	.361*	.314*
Dur. max.	−.169	−.117	−.151	−.402**	.010	−.273*

"FP" = Free Play; "SF" = Still Face challenge; "RSA" = respiratory sinus arrhythmia; "HP" = heart period; "Fund." = fundamental frequency; "BW" = 50% bandwidth; "Var." = variance; "Dur." = duration.

* $p < 0.05$, Pearson correlation (2 tailed).

** $p < 0.01$, Pearson correlation (2 tailed).

2.6. Statistical analysis

Sex differences were assessed using both unpaired t - and U -tests. Correlations were calculated to evaluate the relationships between vocalization features and measures of autonomic state (i.e., heart period, RSA). Correlations were also calculated to evaluate the relationships among the individual acoustic variables. Relationships among the acoustic variables were further explored by a rotated factor analysis (varimax with Kaiser normalization) using principal component analysis as the extraction method. Correlations were calculated to assess the relationship between each of the factors and the autonomic measures. All statistical analyses were performed using SPSS 18.0 (SPSS, Inc., Chicago, IL).

3. Results

The descriptive features of the prosodic and autonomic variables are reported in Table 2. While sex differences were largely not present, significant differences were observed between males and females for minimum variance and duration maximum ($p < 0.05$; U -test). Significant correlations were observed between several acoustic features of the infant vocalizations and the indices of autonomic state are reported in Table 3. As described in Table 3, shorter heart period (i.e., faster heart rate) during the Still Face procedure was observed in infants with higher maximum fundamental frequency, greater maximum bandwidth and variability, and longer maximum duration cries. The upper limits of these extreme values for acoustic features represent atypical cries (e.g., long wailing shrieks), which were associated with faster heart rate. Similarly, infants with the greatest suppression from Free Play to Still Face of heart period (i.e., heart rate increase) and RSA (i.e., cardiac vagal tone withdrawal) had higher maximum values of all the acoustic variables. Furthermore, longer heart period (i.e., slower heart rate) was associated with higher minimum values of 50% bandwidth and variance. RSA suppression showed the same pattern: the change in RSA exhibited a significant positive correlation with the variance and duration minimum (Table 3).

Overall, the convergence of the positive and negative relationships with autonomic state is best described by a U-type relationship, whereby the minimal and maximal values for acoustic parameters that are closer to the median (i.e., higher minima values and lower maxima values), represent an expansion of prosody and a richness of vocal features. Individuals, who produce vocalizations with fewer outliers, tend to have slower heart rates, less heart rate acceleration and less vagal withdrawal (i.e., decrease in RSA) in response to the Still Face paradigm.

Table 4

Correlation between acoustic variables for infant vocalizations measured during the Still Face challenge.

	Fund. Med.	Fund. Min.	Fund. Max.	BW Med.	BW Min.	BW Max.	Var. Med.	Var. Min.	Var. Max.	Dur. Med.	Dur. Min.	Dur. Max.
Fund. med.	1	.466**	.475**	-.209	-.248*	.270*	.329**	-.017	.302*	.261*	.043	.237
Fund. min.	.466**	1	-.024	-.095	.185	-.298**	-.122	.258*	-.245	.333*	.499**	-.207
Fund. max.	.475**	-.024	1	-.042	-.343**	.528**	.371**	-.167	.537**	.008	-.213	.270*
BW med.	-.209	-.095	-.042	1	.664**	.395**	-.209	-.031	-.182	-.472**	-.151	-.363**
BW min.	-.248*	.185	-.343**	.664**	1	-.054	-.380**	.148	-.479**	-.282*	.099	-.584**
BW max.	.270*	-.298**	.528**	.395**	-.054	1	.232	-.250	.422**	-.220	-.382**	.144
Var. med.	.329**	-.122	.371**	-.209	-.380**	.232	1	.519**	.804**	.475**	.101	.739**
Var. min.	-.017	.258*	-.167	-.031	.148	-.250	.519**	1	.132	.434**	.453**	.133
Var. max.	.302*	-.245	.537**	-.182	-.479**	.422**	.804**	.132	1	.311*	-.081	.714**
Dur. med.	.261*	.333*	.008	-.472**	-.282*	-.220	.475**	.434**	.311*	1	.789**	.581**
Dur. min.	.043	.499**	-.213	-.151	.099	-.382**	.101	.453**	-.081	.789**	1	.098
Dur. max.	.237	-.207	.270*	-.363**	-.584**	.144	.739**	.133	.714**	.581**	.098	1

"Fund." = fundamental frequency; "BW" = 50% bandwidth; "Var." = variance; "Dur." = duration.

* $p < 0.05$, Pearson correlation (2 tailed).** $p < 0.01$, Pearson correlation (2 tailed).

Thus, as heart rate increases during "mobilization" of the ANS, such as during the Still Face challenge, an overall higher pitch and flat affect is heard in the vocalization. The shift toward a "monotone" sound with decreasing RSA is further characterized by brief "choppy" vocalizations with a narrowed pitch. Acoustic minima values that did not significantly correlate also showed a positive trend for the autonomic measures obtained following the Still Face challenge.

Individual differences in RSA during the Still Face were not related to the acoustic properties of the infant vocalizations. RSA and heart period during Free Play 1 were not correlated with any of the acoustic features, suggesting that initial state is not predictive of vocalization features during the stressor.

There were significant covariations across several of the acoustic variables. Fundamental frequency, in particular, correlated with all other variables (i.e., 50% bandwidth, variance, and duration) (Table 4). Specifically, the *maximum* values of 50% bandwidth, variance, and duration all demonstrated a shared directionality in their covariation with *maximum* fundamental frequency (i.e., higher pitch cries). Moreover, the *median* value of each acoustic measure consistently demonstrated a shared directionality and covariation with its *minimum* value. For example, subjects with higher median fundamental frequency cries during the challenge also had higher minimum fundamental frequency cries. However, shared directionality was not consistent across all acoustic indices, as an inverse relationship was observed among several acoustic variables as well (e.g., 50% bandwidth maximum in contrast to duration minimum) (Table 4).

A rotated (varimax with Kaiser normalization) principal component factor analysis was applied to the acoustic variables to explore potential relationships among the acoustic measures. The analysis identified a

four factor solution with three factors consisting of the "prosodic variables" (i.e., variance, 50% bandwidth, and duration, in addition to the fundamental indices) and a fourth factor comprised of only fundamental frequency measures (Table 5; Fig. 1). The correlation between each of the four factors and the autonomic measures was then assessed, showing significant covariation between the variables associated with prosodic modulation (factors 1–3) and autonomic state. Factor 1 demonstrated a negative covariation with heart period and heart period change. Factor 2 exhibited a positive covariation with changes in both heart period and RSA. Factor 3 positively covaried with the change in heart period. Uniquely, factor 4, based on the level of the fundamental frequency, did not covary with any of the autonomic measures (Table 6). Notably, none of the factors correlated with either heart period or RSA during Free Play. Rather, as heart rate accelerates during the Still Face challenge, certain acoustic features are differentially recruited to varying degrees to comprise the overall expansion of prosody.

The correlations between the factor scores and the autonomic variables paralleled the correlations between the individual autonomic variables and the specific acoustic features (Table 2). As illustrated in Table 6, Factor 1, which was influenced most strongly by the maximum indices, was negatively correlated with the autonomic variables. In contrast, Factors 2 and 3, which were influenced most strongly by the minimum indices, were positively correlated with autonomic variables. Although the fundamental frequency variables were correlated with the autonomic variables (see Table 3), Factor 4, which was primarily defined by fundamental frequency, was not significantly correlated with any autonomic variable. Overall, the factor analysis further supported the association between decreased acoustic modulation (i.e., restricted prosody) with reduced RSA and heart rate increase, both indicators of cardiac vagal tone.

4. Discussion

While the convergence of RSA and fundamental frequency has been reported in newborns following circumcision [29], prior to this study, modulation of vocalization features in relation to autonomic state had not, to our knowledge, been investigated in human infants or other mammals. Through the application of a novel computation-based tool, we have detected and quantified features of prosody and have demonstrated the convergence between these acoustic features and measures autonomic state. Thus, the study provides initial confirmation that features of infant vocal prosody may provide a sensitive index of autonomic activity in human infants.

The hypotheses linking vocalizations to autonomic state were based on a knowledge of the common neurophysiological mechanisms regulating vocalization and the autonomic nervous system [2,5,43]. In particular, the same brainstem structures are involved in both

Table 5

Relationships among the covariation of the acoustic measures (Table 4), as revealed by rotated factor analysis (varimax with Kaiser normalization) using principal component analysis as the extraction method. Variables exhibiting minimal (<0.3) contribution are excluded below.

	Factor 1	Factor 2	Factor 3	Factor 4
Fund. med.				.912
Fund. min.	-.347	.559		.674
Fund. max.	.476	-.380		.580
BW med.			.832	
BW min.	-.314		.874	
BW max.	.490	-.605		
Var. med.	.927			
Var. min.	.339	.637	.543	
Var. max.	.947			
Dur. med.	.416	.784		
Dur. min.		.885		
Dur. max.	.835		-.414	

"Fund." = fundamental frequency; "BW" = 50% bandwidth; "Var." = variance; "Dur." = duration.

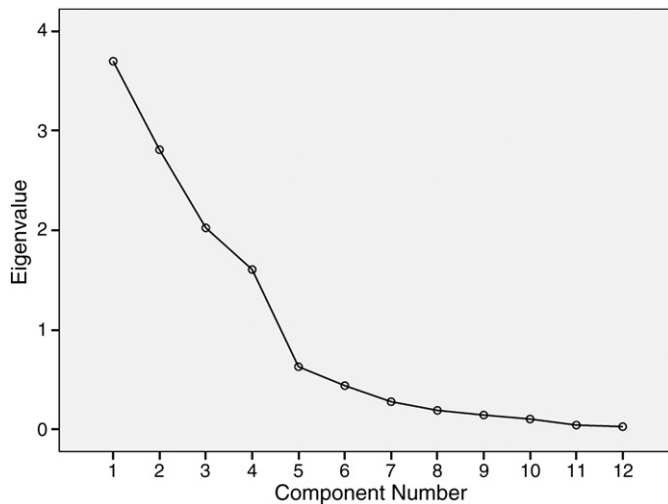


Fig. 1. Components extracted by principal component factor analysis (varimax rotation with Kaiser normalization). Note the presence of 4 factors representing the primary contributions to the acoustic measures' covariation (Table 4).

the regulation of heart rate via the vagus (i.e., cardiac vagal tone) and vocalizations via the laryngeal and pharyngeal muscles (both efferents of the supradiaphragmatic vagus). Consistent with the Polyvagal Theory, the role of these neural pathways as functional neuro-modulators would have effects on both autonomic state (i.e., vagal regulation of the heart) and the acoustic features of vocalizations (i.e., prosody). The findings described here further support the phylogenetic hierarchy proposed by the Polyvagal Theory and suggest that paralleling the phylogenetic shifts in the autonomic nervous system an increased interplay between the vagal regulation of the heart and the striated muscles of the face and head evolved to promote a more coordinated biobehavioral response to challenges [2,3,40]. Specifically, the mammalian vagus is neuroanatomically linked to the cranial nerves that regulate social engagement through processes manifested as facial expression, hearing, and, of particular interest to the current study, vocalization [2,5]. Thus, the coevolution of the physiological processes, underpinning both the ANS and vocalization mechanisms as well as their shared innervations from the vagus, underscores their covariation.

Fundamental frequency level, 50% bandwidth, variance, and duration correlated to one or more measures of autonomic state (i.e., heart period during Still Face and/or change in heart period or RSA from Free Play 1 to Still Face; Table 3). In particular, the *negative* correlations between the *maximum*, and the *positive* correlations between the *minimum*, of these variables and autonomic state illustrates that a more "mobilized" autonomic nervous system characterized by features of vagal withdrawal was associated with a dampening of vocal prosody. This convergence with autonomic state was characterized by a U-type relationship, in which the tendency of the acoustic parameters' extreme to be closer to the median observed values represents an expansion of prosody.

The infant vocalizations investigated in this study were from a subset of infant participants in a still face procedure and represented only infants who that vocalized. The vocalizations studied included a wide range of acoustic variability, ranging from social vocalizations to the high pitch cries typically observed in severely stressed infants. The acoustic maxima obtained represent the extreme of infant vocalizations, drawing from the loudest, most intense cries. Thus, while the acoustic modulation of these cries is restricted with vagal withdrawal, the values of the acoustic maxima (i.e., acoustic energy) are nevertheless higher. Furthermore, heart period and RSA during Free Play 1 were not correlated with any of the acoustic features (32), suggesting that initial state is not predictive of vocalization features. Conversely, of the acoustic features which did covary during the Still Face challenge, it was the more "prosodic variables" that contributed most robustly to the covariation, as opposed to the level of the fundamental frequency (Table 5).

The positive relationship that exists between the height of the larynx muscle and fundamental frequency, with frequency changing relative to changing larynx height (i.e., pitch modulation) [48,49], is perhaps the most basic (and robust) regulation of vocal prosody. In particular, this rapid change in fundamental frequency during a vocalization represents a modulation of neural tone to the laryngeal muscles. With less vagal output corresponding to the contraction of the cricothyroid muscle, there is a functional increase in pitch as the vocal cord is elongated. Conversely, increased vagal output corresponds to contraction of the anatomically antagonistic thyroarytenoid muscle, which serves to lower pitch. The observed negative correlation between heart period and fundamental frequency level is consistent with this neurophysiological model.

Correlations among the individual acoustic measures were assessed and indicated a common covariation among several indices (Table 4). For example, the *median* value of each acoustic measure consistently demonstrated a shared directionality and covariation with its *minimum* value. In line with the shared, positive relationship between laryngeal neural tone and fundamental frequency [48,49], a similar relationship can be expected for the other prosodic features inherently corresponding to the fundamental. Notably, the *maximum* values of 50% bandwidth, variance, and duration all demonstrated a shared directionality in their covariation with *maximum* fundamental frequency. Similarly, the *median* values of variance and duration also exhibited a positive relationship to the *median* fundamental frequency. However, this relationship was inconsistent across other acoustic measures, with several indices exhibiting an inverse relationship to fundamental frequency (Table 4). For example, the fundamental median, while demonstrating a common positive relationship with all other significant measures, expressed the opposite relationship with the 50% bandwidth minimum. Thus, the role of other cortical and subcortical regions on vocal prosody should also be considered [47,58–60], underscoring the observation that mammals can incorporate voluntary influences to convey different aspects of visceral state.

Whereas laryngeal motor control during vocalization is also influenced by interactions with other cortical and subcortical regions, especially those principally involved in the voluntary control of vocalization such as the laryngeal motor cortex [58,59] and anterior

Table 6
Correlations between acoustic factors (obtained via factor analysis; Table 5) and autonomic variables. RSA and heart period changes based on difference between measures obtained from Free Play 1 to the Still Face procedure (i.e., SF–FP1).

	RSA during FP1	HP during FP1	RSA during SF	HP during SF	RSA change	HP change
Factor 1	.014	–.096	–.146	–.496**	–.151	–.348*
Factor 2	–.171	–.141	.280	.264	.471**	.381*
Factor 3	.057	–.101	.014	.293	–.010	.346*
Factor 4	–.017	.076	.003	–.177	–.009	–.250

"FP" = Free Play; "SF" = Still Face challenge; "RSA" = respiratory sinus arrhythmia; "HP" = heart period.

* $p < 0.05$, Pearson correlation (2 tailed).

** $p < 0.01$, Pearson correlation (2 tailed).

cingulate cortex [47,60], the deviations from our proposed model by several measures are not surprising, such as the *positive* covariation of duration median (relative to the *negatively* covarying fundamental frequency and variance medians; Table 3). Through the evolution of voluntary speech, the influence of the laryngeal motor cortex gained prominence as an additional control center of vocalization in humans. This development is evidenced in the laryngeal motor cortex's role in nonhuman primates, in which only indirect connections to the laryngeal motor neurons are present [58,59]. Overall, the cortical-mediated control of vocalizations in addition to vagal regulation further enrich vocal prosody, and represent the use of voluntary influences to communicate different qualities reflecting visceral state.

While the present study demonstrates an overall covariance of vocal prosody with autonomic state, specific acoustic indices could correlate differently in future studies. In particular, different and varying levels of challenges may encourage distinct prosodic features based on what may need to be signaled. Thus, challenges that necessitate the communication of differential signals or visceral states (e.g., pain versus fear) may employ different prosodic features to convey this information.

5. Conclusions

An automated tool recently developed by our laboratory was applied to objectively characterize several features of vocal prosody, further demonstrating the utility of these prosodic features as a sensitive index of autonomic activity in human infants. As predicted by the Polyvagal Theory, the observations of this study support the hypothesis that a common neurophysiological process regulates vocalization prosody and the autonomic nervous system. Specifically, heart period during the Still Face challenge, as well as changes in heart period and RSA, were negatively correlated with the maximum acoustic values, and positively correlate with the minimum acoustic values. These correlations demonstrate an association between shorter heart period and the dampening of the modulation of acoustic features that are perceived as prosody. These data parallel clinical observations in which higher pitch (i.e., fundamental frequency) and less variability in acoustic features of vocalizations (i.e., less vocal prosody, commonly describes as “monotone”) convey to the listener that the autonomic state of the crying infant is more “mobilized” and characterized by features of vagal withdrawal (i.e., reduced RSA) and perhaps sympathetic excitation (e.g., increased heart rate in excess of that predicted by the synchronous reduction in RSA).

Acknowledgments

The project described was conducted as part of the doctoral dissertation of AMS. This study was supported, in part, by Award Number R01HD053570 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- [1] Porges SW, Furman SA. The early development of the autonomic nervous system provides a neural platform for social behavior: a polyvagal perspective. *Infant Child Dev* 2011;20(1):106–18.
- [2] Porges SW. The polyvagal theory: phylogenetic substrates of a social nervous system. *Int J Psychophysiol* 2001;42(2):123–46.
- [3] Porges SW. The polyvagal theory: new insights into adaptive reactions of the autonomic nervous system. *Cleve Clin J Med* 2009;76(Suppl. 2):S86–90.
- [4] Porges SW. Social engagement and attachment: a phylogenetic perspective. *Ann N Y Acad Sci* 2003;1008:31–47.
- [5] Porges SW, Lewis GF. The polyvagal hypothesis: common mechanisms mediating autonomic regulation, vocalizations and listening. In: Brudzynski SM, editor. *Handbook of mammalian vocalization: an integrative neuroscience approach*. London: Academic Press; 2009. p. 255–64.
- [6] Brudzynski SM. Communication of adult rats by ultrasonic vocalization: biological, sociobiological, and neuroscience approaches. *ILAR J* 2009;50(1):43–50.
- [7] Hammerschmidt K, Fischer J. The vocal repertoire of Barbary macaques: a quantitative analysis of a graded signal system. *Ethology* 1998;104(3):203–16.
- [8] LaGasse LL, Neal AR, Lester BM. Assessment of infant cry: acoustic cry analysis and parental perception. *Ment Retard Dev Disabil Res Rev* 2005;11(1):83–93.
- [9] Wood RM, Gustafson GE. Infant crying and adults' anticipated caregiving responses: acoustic and contextual influences. *Child Dev* 2001;72(5):1287–300.
- [10] Lounsbury ML, Bates JE. The cries of infants of differing levels of perceived temperamental difficulty: acoustic properties and effects on listeners. *Child Dev* 1982;53(3):677–86.
- [11] Lester BM, Boukydis CF, Garcia-Coll CT, Peucker M, McGrath MM, Vohr BR, et al. Developmental outcome as a function of the goodness of fit between the infant's cry characteristics and the mother's perception of her infant's cry. *Pediatrics* 1995;95(4):516–21.
- [12] Porter FL, Miller RH, Marshall RE. Neonatal pain cries: effect of circumcision on acoustic features and perceived urgency. *Child Dev* 1986;57(3):790–802.
- [13] Zeskind PS, Lester BM. Acoustic features and auditory perceptions of the cries of newborns with prenatal and perinatal complications. *Child Dev* 1978;49(3):580–9.
- [14] Corwin MJ, Lester BM, Sepkoski C, Peucker M, Kayne H, Golub HL. Newborn acoustic cry characteristics of infants subsequently dying of sudden infant death syndrome. *Pediatrics* 1995;96(1 Pt 1):73–7.
- [15] Koivisto M. Cry analysis in infants with Rh haemolytic disease. *Acta Paediatr Scand* 1987;335:1–73.
- [16] Lehr VT, Zeskind PS, Ofenstein JP, Cepeda E, Warrier I, Aranda JV. Neonatal facial coding system scores and spectral characteristics of infant crying during newborn circumcision. *Clin J Pain* 2007;23(5):417–24.
- [17] Michelsson K, Sirvio P, Wasz-Hockert O. Pain cry in full-term asphyxiated newborn infants correlated with late findings. *Acta Paediatr Scand* 1977;66(5):611–6.
- [18] Quick ZL, Robb MP, Woodward LJ. Acoustic cry characteristics of infants exposed to methadone during pregnancy. *Acta Paediatr* 2009;98(1):74–9.
- [19] Rautava L, Lempiäinen A, Ojala S, Parkkola R, Rikala H, Lapinleimu H, et al. Acoustic quality of cry in very-low-birth-weight infants at the age of 1 1/2 years. *Early Hum Dev* 2007;83(1):5–12.
- [20] Douglas P, Hill P. Managing infants who cry excessively in the first few months of life. *BMJ* 2011;343:d7772.
- [21] Evanoo G. Infant crying: a clinical conundrum. *J Pediatr Health Care* 2007;21(5):333–8.
- [22] Miller J, Newell D. Prognostic significance of subgroup classification for infant patients with crying disorders: a prospective cohort study. *J Can Chiropr Assoc* 2012;56(1):40–8.
- [23] Kim JS. Excessive crying: behavioral and emotional regulation disorder in infancy. *Korean J Pediatr* 2011;54(6):229–33.
- [24] Zeskind PS, Barr RG. Acoustic characteristics of naturally occurring cries of infants with “colic”. *Child Dev* 1997;68(3):394–403.
- [25] St James-Roberts I. What is distinct about infants' “colic” cries? *Arch Dis Child* 1999;80(1):56–61 [discussion 62].
- [26] Soltis J. The signal functions of early infant crying. *Behav Brain Sci* 2004;27(4):443–58 [discussion 459–90].
- [27] Robb MP, Crowell DH, Dunn-Rankin P, Tinsley C. Cry features in siblings of SIDS. *Acta Paediatr* 2007;96(10):1404–8.
- [28] (a) Lester BM. Spectrum analysis of the cry sounds of well-nourished and malnourished infants. *Child Dev* 1976;47(1):237–41; (b) Lester BM, Tronick EZ, LaGasse L, Seifer R, Bauer CR, Shankaran S, et al. The maternal lifestyle study: effects of substance exposure during pregnancy on neurodevelopmental outcome in 1-month-old infants. *Pediatrics* 2002;110(6):1182–92; (c) Zeskind PS, Marshall TR, Goff DM. Cry threshold predicts regulatory disorder in newborn infants. *J Pediatr Psychol* 1996;21(6):803–19.
- [29] Porter FL, Porges SW, Marshall RE. Newborn pain cries and vagal tone: parallel changes in response to circumcision. *Child Dev* 1988;59(2):495–505.
- [30] Scattoni ML, Gandhi SU, Ricceri L, Crawley JN. Unusual repertoire of vocalizations in the BTBR T + tf/J mouse model of autism. *PLoS One* 2008;3(8):e3067.
- [31] Fischer J, Hammerschmidt K. Ultrasonic vocalizations in mouse models for speech and socio-cognitive disorders: insights into the evolution of vocal communication. *Genes Brain Behav* 2011;10(1):17–27.
- [32] Wöhr M, Schwarting RKW. Rodent ultrasonic communication and its relevance for models of neuropsychiatric disorders. *e-Neuroforum* 2010;1(4):71–80.
- [33] Crockford C, Wittig RM, Mundry R, Zuberbühler K. Wild chimpanzees inform ignorant group members of danger. *Curr Biol* 2012;22(2):142–6.
- [34] Kim EJ, Kim ES, Covey E, Kim JJ. Social transmission of fear in rats: the role of 22-kHz ultrasonic distress vocalization. *PLoS One* 2010;5(12):e15077.
- [35] Taracha E, Hamed A, Krząśnik P, Lehner M, Skórzewska A, Płaźnik A, et al. Inter-individual diversity and intra-individual stability of amphetamine-induced sensitization of frequency-modulated 50-kHz vocalization in Sprague–Dawley rats. *Psychopharmacology (Berl)* 2012;222(4):619–32.
- [36] Calvino B, Besson JM, Boehrer A, Depaulis A. Ultrasonic vocalization (22–28 kHz) in a model of chronic pain, the arthritic rat: effects of analgesic drugs. *Neuroreport* 1996;7(2):581–4.
- [37] Scherer KR. Vocal affect expression: a review and a model for future research. *Psychol Bull* 1986;99(2):143–65.
- [38] Garcia-Toro M, Talavera JA, Saiz-Ruiz J, Gonzalez A. Prosody impairment in depression measured through acoustic analysis. *J Nerv Ment Dis* 2000;188(12):824–9.
- [39] Porges SW. Vagal tone: a physiologic marker of stress vulnerability. *Pediatrics* 1992;90(3 Pt 2):498–504.
- [40] Porges SW. Cardiac vagal tone: a physiological index of stress. *Neurosci Biobehav Rev* 1995;19(2):225–33.
- [41] Larson SK, Porges SW. The ontogeny of heart period patterning in the rat. *Dev Psychobiol* 1982;15(6):519–28.

- [42] Field T, Diego M. Vagal activity, early growth and emotional development. *Infant Behav Dev* 2008;31(3):361–73.
- [43] Porges SW. The Polyvagal Theory: phylogenetic contributions to social behavior. *Physiol Behav* 2003;79(3):503–13.
- [44] Bieger D, Hopkins DA. Viscerotopic representation of the upper alimentary tract in the medulla oblongata in the rat: the nucleus ambiguus. *J Comp Neurol* 1987;262(4):546–62.
- [45] Ayres JG, Gabbott PLA. Vocal cord dysfunction and laryngeal hyperresponsiveness: a function of altered autonomic balance? *Thorax* 2002;57(4):284–5.
- [46] Newman JD. Evolution of the communication brain in control of mammalian vocalizations. In: Brudzynski SM, editor. *Handbook of mammalian vocalization: an integrative neuroscience approach*. London: Academic Press; 2009. p. 23–8.
- [47] Jürgens U. The neural control of vocalization in mammals: a review. *J Voice* 2009;23(1):1–10.
- [48] Hirano M, Ohala J, Vennard W. The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. *J Speech Hear Res* 1969;12(3):616–28.
- [49] Ohala JJ. How is pitch lowered? *J Acoust Soc Am* 1972;52(1A):124.
- [50] Hoh JFY. Laryngeal muscles as highly specialized organs in airway protection, respiration and phonation. In: Brudzynski SM, editor. *Handbook of mammalian vocalization: an integrative neuroscience approach*. London: Academic Press; 2009. p. 13–22.
- [51] Eckberg DL. The human respiratory gate. *J Physiol* 2003;548(Pt 2):339–52.
- [52] Porges SW. Orienting in a defensive world: mammalian modifications of our evolutionary heritage. A Polyvagal Theory. *Psychophysiology* 1995;32(4):301–18.
- [53] Andersen MC, Kunze DL. Nucleus tractus solitarius — gateway to neural circulatory control. *Annu Rev Physiol* 1994;56(1):93–116.
- [54] Weinberg MK, Tronick EZ. Infant affective reactions to the resumption of maternal interaction after the still-face. *Child Dev* 1996;67(3):905–14.
- [55] Porges SW, Bohrer RE. The analyses of periodic processes in psycho-physiological research. In: Cacioppo JT, Tassinari LG, editors. *Principles of psychophysiology: physical, social and inferential elements*. New York: Cambridge University Press; 1990. p. 708–53.
- [56] Porges SW. Method and apparatus for evaluating rhythmic oscillations in aperiodic physiological response systems; 1985.
- [57] Lewis GF, Furman SA, McCool MF, Porges SW. Statistical strategies to quantify respiratory sinus arrhythmia: are commonly used metrics equivalent? *Biol Psychol* 2012;89(2):349–64.
- [58] Simonyan K, Horwitz B. Laryngeal motor cortex and control of speech in humans. *Neuroscientist* 2011;17(2):197–208.
- [59] Brown S, Ngan E, Liotti M. A larynx area in the human motor cortex. *Cereb Cortex* 2008;18(4):837–45.
- [60] von Cramon D, Jürgens U. The anterior cingulate cortex and the phonatory control in monkey and man. *Neurosci Biobehav Rev* 1983;7(3):423–5.