

**Provided for non-commercial research and educational use only.
Not for reproduction, distribution or commercial use.**

This chapter was originally published in the book *Handbook of Behavioral Neuroscience, Vol. 25* published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who know you, and providing a copy to your institution's administrator.

Handbook of Behavioral Neuroscience

Series Editor: Joseph P. Huston

Handbook of Ultrasonic Vocalization

A Window into the Emotional Brain



Edited by Stefan M. Brudzynski

Volume 25



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<https://www.elsevier.com/permissions>

From Jacek Kolacz, Gregory F. Lewis and Stephen W. Porges, The Integration of Vocal Communication and Biobehavioral State Regulation in Mammals: A Polyvagal Hypothesis. In: Stefan M. Brudzynski, editor, *Handbook of Behavioral Neuroscience, Vol. 25*, Amsterdam: Academic Press, 2018, pp. 23-34.

ISBN: 978-0-12-809600-0

© Copyright 2018 Stephen Porges

Academic Press

The Integration of Vocal Communication and Biobehavioral State Regulation in Mammals: A Polyvagal Hypothesis

Jacek Kolacz¹, Gregory F. Lewis^{1,2}, Stephen W. Porges^{1,3}

¹Traumatic Stress Research Consortium, Kinsey Institute, Indiana University, Bloomington, IN, United States

²Intelligent Systems Engineering, Indiana University, Bloomington, IN, United States

³Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

I INTRODUCTION

Vocalizations are a key part of the social lives of mammals, facilitating the coregulation of internal states related to threat response and affiliative contact between conspecifics. This chapter, informed by the Polyvagal Theory (Porges, 1995; 2001; 2007; 2011) applies an evolutionary approach to the neurophysiological and communicative functions involved in mammalian vocalizations. We posit that the encoding of safety and threat cues in mammalian vocalizations is regulated through the a neurophysiologically integrated social engagement system, which involves pathways in several cranial nerves that regulate muscles of the face and head (via special visceral efferent pathways) and the physiological state (via vagal pathways). Using this perspective, we propose a model of mammalian vocal communication wherein emotional states are functionally coordinated with processes that modulate acoustic features of vocalizations and filter acoustic information from the environment. This model posits that the signal conduction properties of the middle ear place constraints on the frequency band of perceptual advantage, a frequency range in which conspecifics communicate with one another (see also Porges & Lewis, 2010), and that this band determines the specific frequency ranges within which cues of safety and threat are communicated. These frequency constraints and effects on listeners' internal states directly correspond to the regulation of vocal acoustic features by laryngeal and pharyngeal neural control, which are neurally integrated in the brainstem with regulation of the vocalizer's emotional states.

This chapter first provides an overview of the mammalian neural social engagement system and its evolutionary origins. We then describe its function in integrating middle ear neural regulation with physiological states. Finally, we present an overview of the common mechanisms that regulate the acoustic features of vocalizations using vocal fold vibration and rodent "whistle-like" vocalizations. We include demonstrations of the regulation of vocal production and listening with examples from human infant cries, whistle vocalizations of prairie voles (*Microtus ochrogaster*), and elephant calls, representing mammals that span a wide range of middle ear bone sizes. The chapter closes with a general mammalian model of vocal production and reception drawn from the reviewed literature.

II EVOLUTION AND ORGANIZATION OF THE MAMMALIAN SOCIAL ENGAGEMENT SYSTEM

The Polyvagal Theory proposes that phylogenetic changes in the neural regulation of vertebrate autonomic systems enabled in mammals a face-heart connection that facilitates bidirectional state regulation between conspecifics (Porges, 1995; 2001; 2007; 2011). Comparative neuroanatomy documents that the neural regulation of the heart shifted during vertebrate evolution toward more rapid and efficient neural regulation to meet fluctuating metabolic demands. This shift was characterized by a progression from largely endocrine control in primitive nervous systems to the addition of the faster

unmyelinated nerves to the most recent evolutionary innovation of rapidly responding myelinated nerves. In mammals, the regulation of the heart is linked to other supradiaphragmatic structures through the ventral vagal complex, a system of neural projections that includes motor source subnuclei originating in the nucleus ambiguus of the brainstem. This nucleus is positioned ventrally to the dorsal motor nucleus of the vagus, which provides source subnuclei for the dorsal vagal complex, a largely unmyelinated vagal system that predominantly innervates subdiaphragmatic organs.

The evolutionary emergence of myelinated vagal pathways co-occurred with functional and anatomical changes in the neural regulation of the muscles and structures of the head and face. This included the detachment of the middle ear bones from the jaw and brainstem integration of the source nuclei regulating myelinated cardiac vagal fibers with those regulating the striated muscles of the face and head through special visceral efferent pathways. These neural changes permitted the integration of supradiaphragmatic body regulation (e.g., heart and bronchi) with that of the facial and head muscles that were previously used largely for basic survival needs such as defensive and aggressive biting, ingestion, chewing, and airway protection. This integration links the regulation of the heart and bronchi with the muscles that provide expressivity around the eyes, control head turning for social gesturing, promote efficient extraction of conspecific vocalizations from background sounds by the middle ear muscles, and regulate the acoustic features of emitted vocalizations via the larynx and pharynx.

This integrated social engagement system is the most phylogenetically recent state-regulation system to emerge in vertebrates, reflecting selective pressure for cooperative behaviors in mammals. Unlike mammals' reptilian ancestors, whose nervous systems were adapted for survival in dangerous and life-threatening contexts, the mammalian social engagement system evolved to promote social interactions and bonding with conspecifics in safe contexts. The social engagement system promotes calm, restorative states that support affiliative, prosocial behavior. Control via this system reduces blood pressure, decreases heart rate, reduces metabolic demands, and optimizes homeostatic processes that support health, growth, and restoration.

Polyvagal Theory identifies three primary neurophysiological states. These states are modulated by neuroception, a nonconscious neural process that evaluates risk in the environment using signals of safety and threat that have been shaped by evolution (see [Porges, 2004](#), [2015](#)). Neuroception of safety promotes state regulation by the social engagement system while danger or life-threatening cues promote the rapid and graded functional withdrawal of the myelinated vagal system, allowing older survival-based systems to activate in

order to meet environmental challenges. In the case of such challenges, state regulation is proposed to follow the principles of Jacksonian dissolution, with phylogenetically older, survival-oriented systems becoming recruited when more recent circuits fail to enable the organism to navigate to safety. Mobilization is facilitated by the sympathetic nervous system while immobilization, such as death feigning and vasovagal syncope (fainting), is regulated by the unmyelinated dorsal vagal complex that is shared among most vertebrates.

This integration of state regulation with social communication provides a biobehavioral means for rapid social propagation of danger and threat signals as well as promotion of affiliative behaviors. Auditory signals are one key aspect of such communication. The neural modulation of vocal features and their neurally mediated effects on the physiological states of conspecifics provide a channel for the coregulation of states in dyads or groups. In group-living mammals, this acoustic channel allows for rapid mobilization in response to acoustic cues of threat (facilitating shared vigilance for danger) or need for conspecific assistance. In contrast, neurally modulated acoustic signals can be used to promote affiliative and bonding behavior through coregulation of states related to safety (see the concept of ethotransmission in Chapter 10).

III MAMMALIAN LISTENING

A The Frequency Band of Perceptual Advantage

The hearing range of individual mammalian animals varies widely, with smaller animals having higher sensitivity for relatively high frequencies and larger animals for low frequencies ([Heffner & Heffner, 2007](#)). The hearing range for a given mammalian species is primarily dependent on the physics of the structures involved in hearing in the middle and inner ear. Vocalizations in ranges above the upper human hearing limit—about 20,000 Hz—are termed “ultrasonic” while those below the human hearing limit—about 20 Hz—are termed “infrasonic.” Such vocalizations outside human perceptual limits are common in the high-frequency calls of small mammals like rodents and ethotransmission low-frequency calls of large mammals like elephants. The perceptual abilities of rodents are more sensitive at higher frequencies, with most species-typical upper hearing thresholds above 50 kHz ([Heffner & Heffner, 2007](#); [Lange, Stalleicken, & Burda, 2004](#)). Thus, rodent high-frequency calls occur within a frequency range easily perceived by their auditory system and help evade auditory detection by larger predator mammals. Given our comparative approach, this chapter does not employ the human-centric “ultrasonic” distinction but rather uses a

relative approach to vocalization production and reception based on the hearing ranges of individual mammalian groups.

Within their absolute hearing ranges, conspecific mammal groups tend to have heightened sensitivity toward frequencies that are optimized near the resonant frequencies of their middle ear anatomy, described as the frequency band of perceptual advantage (Porges & Lewis, 2010). The human middle ear bones have an average resonant frequency around 1000 Hz under air conduction excitation (Silman & Silverman, 1991). Evolutionary pressures may have constrained social affiliative communication to develop in frequencies near this resonant property, making use of the signal boost benefit conferred by the middle ear to facilitate efficient vocal processing in the presence of other environmental sounds. This boost can be seen in the human audiogram shape, which has a bowl-shaped band of perceptual advantage for frequencies between 500 and 4000 Hz. The multiplier of this frequency band of perceptual advantage is roughly symmetrical around the 1000 Hz resonant frequency on a log₁₀ scale (resonant frequency/lower bound = 1.11, upper bound/resonant frequency = 1.20). This auditory boost corresponds with the human index of articulation (Kryter, 1962), the frequency range within which sounds that are critical to language understanding occur. Although speech sounds are produced within a broad frequency range, frequencies occurring below 500 Hz and above 4000 Hz can be filtered out with little loss of intelligibility and, conversely, filtering out frequencies within the 500–4000 Hz range results in vocal signals that have little intelligible quality.

The signal boost provided by middle ear resonance thus likely provided a physical bias on the frequencies within which human affiliative communicative signals emerged. This may reflect a general mammalian evolutionary pressure for matching conspecific vocal production mechanisms with middle ear resonance. In general, smaller middle ear bones tend to have higher resonating frequencies, promoting a higher frequency band of perceptual advantage. The small physical sizes of the earliest mammals provided a perceptual advantage for vocal communication in high-frequency ranges, outside the perceptible range of large reptilian predators. We propose that vocalization behavior within individual mammalian groups coevolved with these receptive frequency range constraints, contributing to the interspecies diversity of animal vocalization frequency ranges. We further posit that safety and threat signals are embedded in the frequency and spectrotemporal modulation characteristics of vocalizations. Signal interpretation by the receiving animal is determined partly by the distribution of the frequency band of perceptual advantage of the listener. We propose that social, nonthreatening cues tend to occur within a mammalian species-typical frequency

band of perceptual advantage. In contrast, vocalizations reflecting danger or threat cues occur at typical high- or low-frequency ranges that are perceptible but outside the heightened sensitivity range provided by the physics of the middle ear.

These interpretations of acoustic features are based on the phylogenetic evolutionary history of vertebrates. At relatively high frequencies, vocalizations represent distress calls and likely signal nearby conspecifics of injury or threat detection. An example can be found in the infant cry, which occurs in a relatively high spectral range of human hearing and produces caregiver responses to reduce infant distress. Low-frequency sounds, toward the low end of a mammal's frequency band of perceptual advantage, similarly have a link with danger cues. Over the course of mammalian evolutionary history, vocalizations in a relatively low-frequency region were likely to be produced by larger predators. Thus, acoustic signals with high energy above or below a mammal's frequency band of perceptual advantage may trigger resource mobilization for fight/flight responses or the behavioral shutdown behaviors that are used for death feigning.

B Neural Modulation of Auditory Sensitivity

While the frequency band of perceptual advantage is determined in mammals by the physics of the middle ear structures, we also propose that this heightened perceptual sensitivity is neurally regulated as part of a feedback loop. The mammalian middle ear (ME) is a highly specialized transducer that couples the atmospheric air pressure changes to the inner ear sensory system, transforming airborne pressure waves into fluid-borne waves within the cochlea. The ME is one in a series of filters along the transmission pathway from the environment to the cognitive processes resulting in the perception of sound. As a filter, it is sensitive to neuromuscular regulation with changes in muscle tension within the middle ear altering the stiffness of the tympanic membrane (e.g., compliance). The middle ear muscles, via cranial nerves, regulate the position of the ossicles and stiffen or loosen the eardrum. When the eardrum is "tightened," higher frequencies are absorbed and transmitted to the inner ear (cochlea) and the energy of lower frequencies is attenuated (i.e., reflected). The auditory nerve (cranial nerve VIII) transmits the acoustic signal from the inner ear to the cortex. Complementing the ascending pathways are descending pathways that regulate the middle ear muscles, which functionally determine the energy (i.e., attenuate, pass, or amplify) of specific frequencies that reach the inner ear.

The ME plays an essential role in compression within the mammalian auditory system. Its function is similar to an automatic gain control that enables a more linear

processing within a restricted range by higher neural circuits (Zwislocki, 2002). The acoustic stapedial reflex (ASR) is an example of an aspect of this automatic gain control. Loud sounds, detected in the cochlea, trigger a bilateral brainstem reflex that contracts the stapedius muscle (a small muscle stabilizing stapes), reducing the transmission of acoustic energy into the cochlea. The attenuation level of acoustic energy transmission to the cochlea mediated by the ASR is frequency dependent (Liberman & Guinan, 1998; Pang & Guinan, 1997). Greater attenuation occurs at lower frequencies (Pang & Guinan, 1997). Above a 1000 Hz cut-off frequency, tension on the human stapedius muscle increases the reception of signals in the human frequency band of perceptual advantage (Lutman & Martin, 1979). Thus, changing the tonic tension of the stapedius muscle would alter the ASR threshold and serve a critical role in modulating the social information received by the auditory system.

The stapedius muscle is innervated by the stapedial branch of the facial nerve (CN VII), which is integrated as part of the ventral vagal complex. The neural pathway responsible for the ASR is well defined, but there is ambiguity about the role of the multiple synaptic projections to the motor neuron pool of the stapedius that suggests modulation by higher cortical structures (Mukerji, Windsor, & Lee, 2010). Descending pathways from the locus coeruleus, inferior colliculus, and superior olivary complex are either directly or indirectly connected to the stapedial motor neuron pool (Brown & Levine, 2008; Rouiller, Capt, Dolivo, & De Ribaupierre, 1989). We suggest that an adaptive feature of mammalian vocal communication would account for this modulation. Thus we expect that middle ear muscle tension, specifically stapedius muscle tension at rest, is regulated as part of the neural regulation of the striated muscles of the face and head via special visceral efferent pathways originating in the ventral vagal complex.

We propose that danger and safety cues, as assessed via neuroception, can mediate the neural tone to the middle ear muscles and alter the ear's transfer function, changing sound frequency responsivity. This system provides a feedback loop that alters auditory sensitivity based on the danger and safety cues embedded in those sounds. Auditory cues of safety promote increases in sensitivity around the peak of the frequency band of perceptual advantage, around which affiliative social vocalization occurs. This serves the adaptive function of increasing efficient extraction of conspecific social vocal cues from environmental noise while reducing relative responsivity to high- and low-frequency cues associated with danger. On the other hand, auditory cues of danger – which can be communicated at frequencies above and below the frequency band of perceptual advantage – decrease efficient processing of affiliative

vocal cues but increase sensitivity to cues at the high and low range. This serves the adaptive function of increasing sensitivity to danger cues, which, as described earlier, occurs at the high and low ranges of auditory frequency bands.

Indeed, there is evidence that regulation of auditory sensitivities around the frequency band of perceptual advantage can be altered by the acoustic features of environmental sounds. Direct evidence comes from an auditory intervention that targeted individuals with autism spectrum disorders (ASD). The symptoms of ASD may serve as a particularly illustrative example of the integration of the ventral vagal complex, as individuals with these disorders are marked by social deficits and particularly high rates of compromised frequency conduction in the human frequency band of perceptual advantage (Rafal, 2013). These ASD-typical characteristics may be linked through the muscles of the face and head that support affiliative social communication and the regulation of the middle ear transfer function that modulates acoustic sensitivity and facilitates the extraction of conspecific voices from environmental noise. Furthermore, the high severity of social deficits in individuals with ASD is related to weak myelinated vagal regulation of the heart (Patriquin, Scarpa, Friedman, & Porges, 2013).

Recent interventions targeting children, adolescents, and young adults with ASD have used dynamically filtered human vocal music processed to amplify features of vocal prosody within the human frequency band of perceptual advantage. This processed audio was successful in increasing respiratory sinus arrhythmia (an index of myelinated vagal activity to the heart), improving auditory processing of social information (Porges et al., 2013), and reducing parent-reported auditory hypersensitivities, which tend to be in the high- and low-frequency range (Porges et al., 2014). This experimental evidence supports a feedback model of the auditory features of sounds as influencing the regulation of the middle ear in concert with the broader social engagement system.

Although, to our knowledge, there are no studies that have examined the role of acoustic cues altering hearing sensitivities in animal models, evidence shows that experimentally induced states – which likely affect the function of the autonomic nervous system – may modulate acoustic responsivity. An experiment wherein an animal is conditioned to anticipate a shock may cause reduced ventral vagal complex regulation and potentially regulate the middle ear for greater sensitivity to auditory threat cues outside the frequency band of perceptual advantage. In one study, mice tested for auditory sensitivity range using a “punishment” shock condition show dampened sensitivity to acoustic stimuli around 10 and 50 kHz, attenuating the “bowl” shape that characterizes the range around peak responsivity on an audiogram, when compared to mice tested using

a “reward” condition (Markl & Ehret, 1973). The changes in auditory sensitivity are consistent with models of the ME transfer function due to decreased stapedius muscle tension in the “shock” condition. In addition to the ventral vagal complex regulation of auditory sensitivity, the neural projections that innervate the larynx and pharynx also provide a physiological mechanism for coordinating the emotional state with expressive vocal acoustic features.

IV MAMMALIAN VOCALIZATIONS

A Anatomy, Function, and Neural Regulation

Key structures involved in generating and modulating acoustic signals in mammals are the larynx, which houses the vocal folds, and the pharynx, the muscle-lined space that sits above the larynx. Variations abound in how these structures are functionally utilized for the production of vocalizations. Here we focus on two functional vocalization types: those induced by vocal fold vibration, as is used to produce human phonated speech and elephant rumbles (Herbst et al., 2012), and the whistle mechanism used by rodents for high-frequency calls. Despite these functionally distinct processes for generating vocalizations, the comparisons of their functional regulation via the ventral vagal complex are supported by their conserved mammalian structure and innervation of the larynx and pharynx via projections from nucleus ambiguus, which also houses the source neurons of myelinated vagal efferent pathways to the heart and lungs.

The larynx, colloquially known as the “voice box,” is a muscular organ that sits above the trachea and is the source of acoustic energy for phonated sounds. This structure is largely conserved across mammals (e.g., Saigusa, 2010). The first primitive larynx emerged as a single constrictor muscle in air-breathing fish to protect the lung inlet from water, its anatomy later adapted for lung volume control (Berke & Long, 2010). The most recent evolutionary exploitation of the larynx has been as a source of vocalization production through the transformation of airflow from the lungs into acoustic energy. The pharynx, the cavity located immediately superior to the larynx, is lined with muscles that alter the shape and movement of this cavity, supporting feeding and respiration in vertebrates and most invertebrates. Though the morphology of the pharynx is more variable among mammals than that of the larynx, the dynamic manipulation of pharyngeal lining plays an important role in shaping the resonating frequencies of signals arising from the larynx in many mammalian vocalizations (Fitch, 2006).

Many acoustic signals of mammalian vocal communication are generated via vibration of the vocal folds,

which line the inner walls of the larynx. Air expired from the lungs induces self-sustained vocal fold vibration, the speed of which results in the fundamental frequency (F_0) of a vocalization that is typically perceived as “pitch.” Vocal fold tension and the size of the glottis – the space between the folds – affect the acoustic qualities of the generated sound. These features are regulated by the movement and length of the vocal folds, which are dynamically controlled via a complex system of muscles that includes vocal fold adductors that decrease the size of the vocal fold opening, a vocal fold abductor that expands the glottis, and a tensor that regulates vocal fold length (see Hof, 2010 and Chapter 4). The combination of these muscles permits highly precise control over the frequency of the acoustic waves generated by the larynx, the resonating frequencies of which are further altered while traveling through the pharynx.

While many mammals use the above-described mechanism of sound produced via vocal fold vibration and pharyngeal modulation of resonating frequencies, much rodent communication between conspecifics occurs via a whistle process. During these vocalizations, the vocal folds are tightly opposed without vibration while air is pushed through a small opening, resulting in acoustic signals marked by nearly pure tones with little harmonic power and fundamental frequencies largely above 20,000 Hz. These vocalizations provide multiple communication benefits for rodents. Their fundamental frequencies occur outside the audible range of many predators and the poor propagation of their short wavelengths limits audibility to a close vicinity.

The generation of fundamental frequencies within this range is not suited for the vocal fold vibration mechanism described earlier, which is constrained by vocal fold tissue strength. In rats, vocal folds undergo rapid physical deterioration when vibrating at rates higher than 6 kHz (Riede, York, Furst, Müller, & Seelecke, 2011). Although it has been proposed that modulation of frequencies produced by the whistle mechanism is largely caused by changes in airflow while tension in the laryngeal muscles remains stable, recent experimental evidence shows that the laryngeal musculature plays the dominant role in frequency modulation of these vocalizations (e.g., Riede, 2013). Thus, the laryngeal muscular modulation mechanism common to both vocal fold vibration-induced vocalizations (e.g., human phonated speech and elephant rumbles) and the rodent whistle method provides a comparative model for examining the neural regulation of vocal acoustic features as part of the social engagement system. Although humans lack a true laryngeal whistle register, rodent whistle-like vocalizations may have been evolutionarily conserved in the vocalizations of some larger mammals such as dogs, whose whines are generated at high frequencies with little or no harmonic structure and likely without vocal fold vibration.

In mammals, visceromotor regulation of the heart and bronchi includes a myelinated vagal pathway arising from the nucleus ambiguus, which also contains source neurons for the multiple cranial nerve pathways that innervate the larynx and pharynx. These projections carry signals to the larynx through the superior laryngeal nerve and recurrent laryngeal nerve branches of the vagus nerve and to the pharynx through the pharyngeal branch of the vagus nerve and the glossopharyngeal (CN IX) nerve. This structural integration with the visceromotor component that regulates the heart and bronchi via the nucleus ambiguus enables the functional coordination of internal states with expressive features of mammalian vocalizations.

The superior laryngeal nerve that branches from the inferior ganglion of the vagus nerve (also known as nodose ganglion) is composed of efferent motor neurons that innervate the cricothyroid (CT) muscle, a vocal fold tensor. In humans, efferent activity along this pathway has direct effects on the expressive features of vocalizations. An *in vivo* blockade of the external superior laryngeal nerve, which is composed of the motor fibers that innervate the CT muscle, reduces fundamental frequency range and increases mean fundamental frequency in adult males (Roy et al., 2009). Injury or dysfunction leads to substantial limiting of pitch range, vocal fold vibratory phase asymmetry, and acoustic aperiodicity (Orestes & Chhetri, 2014). The recurrent laryngeal nerve, the second source of laryngeal innervation that branches from the vagus nerve near the heart, controls the majority of the laryngeal muscles. Its disruption results in profound difficulties with speech production even when nerve effects are unilateral (Crumley, 1994). Neural regulation of the muscles of the pharynx by the ventral vagal complex is largely coordinated through the pharyngeal branch of the vagus nerve, innervating the constrictor muscles that alter the vocal tract's shape. Further control is provided through the glossopharyngeal nerve to the stylopharyngeus muscle, which contributes to the elevation of the larynx and pharynx as well as pharynx dilation.

B Biobehavioral State Regulation and Vocalizations: Evidence From Human Infant Cries and Animal Models

The mammalian integration of laryngeal and pharyngeal innervation with myelinated vagal efferent pathways routed to the heart and lungs provides the foundation for the coregulation of vocalization acoustic features and biobehavioral state. As reviewed earlier, the neural regulation of vocal acoustic features links the physiological regulation of emotional expression between conspecifics in part based on the listener's frequency band of perceptual advantage. Fig. 3.1 depicts

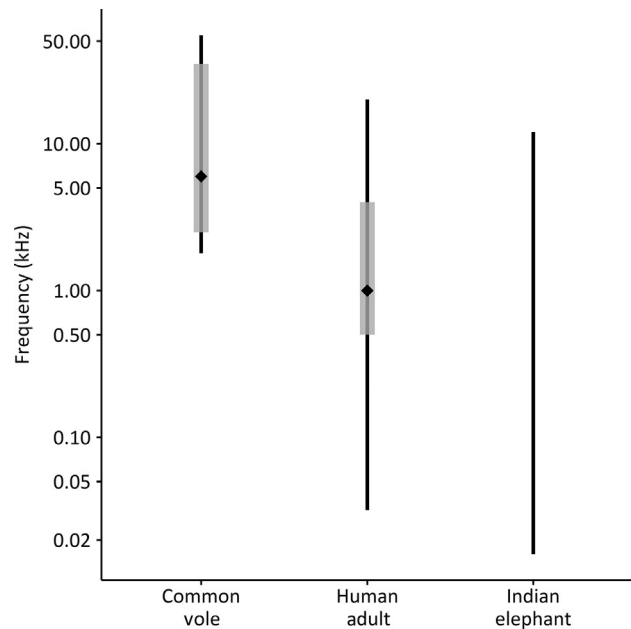


FIG. 3.1 Hearing ranges (thin bars), average middle ear resonant frequency (diamond), and estimated frequency bands of perceptual advantage (thick grey bars) for *Microtus arvalis* (common vole), adult human, and *Elephas maximus* (Indian elephant); common vole resonant frequency and absolute hearing range drawn from Lange et al. (2004); Indian elephant absolute hearing range drawn from Heffner and Heffner (1982).

the hearing ranges and estimated frequency bands of perceptual advantage for an adult human, the Indian elephant (*Elephas maximus*), and the common vole (*Microtus arvalis*). In humans, the frequency band of perceptual advantage spans from about 500–4000 Hz (Porges & Lewis, 2010). Elephant hearing ranges are substantially below those of humans. The estimate provided here is for the *E. maximus* (hearing range: 16–12,000 Hz, lower bound tested at 65 dB; Heffner & Heffner, 1982). To our knowledge, psychoacoustic data for the African elephant (*Loxodonta africana*), which provides the foundation for most elephant vocalization research, is not currently available. However, Indian and African elephants likely have comparable auditory perception due to the anatomical similarity of their auditory structures (Langbauer, 2000; Nummela, 1995). Further, we are not aware of any study that has assessed the resonant frequency of the middle ear in elephants. This lack of information precludes us from estimating a frequency band of perceptual advantage. However, given the large size of the elephant middle ear bones in both African and Indian elephants (Nummela, 1995), this range would be expected to be substantially lower than that of humans. Finally, based on the 6 kHz resonant frequency reported for the common vole (*M. arvalis*; Lange et al., 2004), we estimate the perceptual advantage for these animals may be between approximately 2.5–35 kHz, based on the human multiplier described earlier. We use the *M. arvalis*

(common vole) animal model as a proxy for the auditory response of the anatomically similar *M. ochrogaster* (prairie vole).

In humans, the infant cry is a particularly salient alarm signal and one of the earliest social communication tools available to a neonate. While the cry is a stereotyped behavior that is unquestionably understood to communicate distress, the graded acoustic features of cries can be used as an index of physiological regulation via the ventral vagal complex. A neonate's cries during circumcision, a painful nociceptive stimulus, have the highest fundamental frequency during the most invasive parts of the procedure, which are likely the most distressing and metabolically taxing moments. Infants with the highest cardiac vagal tone (indexed by respiratory sinus arrhythmia) prior to the procedure produce cries with the lowest fundamental frequencies (Porter, Porges, & Marshall, 1988). Negative relations between cardiac vagal tone and infant vocalization fundamental frequencies are also observed in an infant's response to a social challenge (Stewart et al., 2013) and cries in infants who have been born preterm (Shinya, Kawai, Niwa, & Myowa-Yamakoshi, 2016).

In addition to altering mean vocalization frequencies, withdrawal of the ventral vagal complex regulation of the voice may also destabilize musculature control, resulting in rapid unsteady modulation in temporal (rhythmic) and spectral (frequency) domains. Our laboratory has quantified such vocalization modulation using an adapted version of the modulation power spectrum, a two-dimensional Fast Fourier Transformation that decomposes the time-varying acoustic signals of a spectrogram into a two-axis space of spectral and temporal modulation (MPS; Singh & Theunissen, 2003). Preliminary evidence using an MPS-derived measure of modulation depth suggests that infants with higher cardiac vagal tone—reflecting more robust regulation via the ventral vagal complex—have more smoothly modulated cries (Kolacz, 2016; Kolacz, Lewis, Gariépy, & Porges, in preparation; Fig. 3.2).

Human infant cries likely induce ventral vagal complex withdrawal in adult listeners (promoting bodily mobilization of resources to aid the infant) partly due to the relatively high-frequency properties of the signal. A key feature of the stereotyped cry is substantial power above the frequency band of perceptual advantage. The top panels of Fig. 3.3 compare the acoustic signal of a human adult female speaker (top left panel) and a human infant cry (top right panel). Note that infant cries generate more power toward the top and above the frequency band of perceptual advantage (the thick bar on the left vertical line). However, ventral vagal regulation of infant cry fundamental frequency and spectrotemporal modulation provide a mechanism for graded effects on the ventral vagal complex in listeners. Infant cries with high

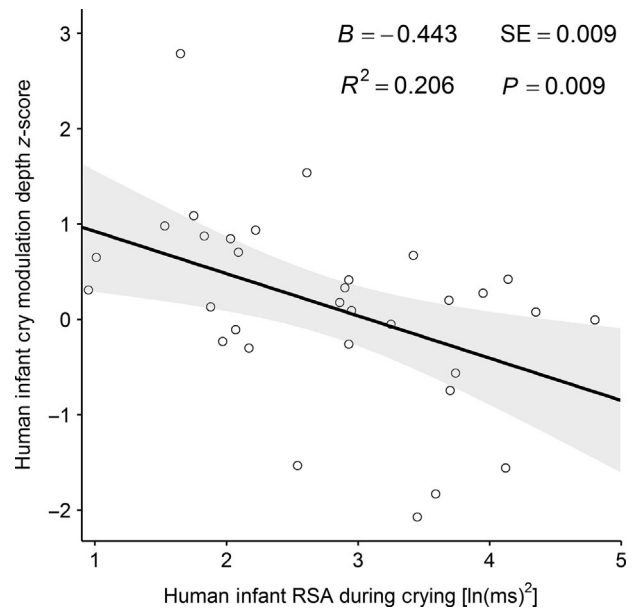


FIG. 3.2 Human infant cry spectrotemporal modulation depth (standardized) as a function of cardiac vagal tone, indexed by respiratory sinus arrhythmia (RSA; method described in Porges, 1985).

fundamental frequencies are perceived as more arousing, urgent, distressing, and aversive in comparison to those whose fundamental frequencies are low (Zeskind & Marshall, 1988). In a study of relative power within frequency bands of infant cries, the amount of power toward the upper range of the human band of perceptual advantage (around 3–4 kHz) was positively related to adult listener ratings of urgency, distress, aversion, and discomfort (Gustafson & Green, 1989). The relatively strong power in this high-frequency range pushes to the upper limit of the human band of perceptual advantage, the zone in which mammalian signaling of distress cues occurs. We posit that these experiences of listener aversion are physiologically induced by triggering autonomic responses to the frequency and modulation characteristics of cries.

The functional withdrawal of the ventral vagal complex can thus aid in mobilizing the adult listener to respond and may couple with sympathetic activation. When compared to typical infant cries, hyperphonated cries—which tend to have higher mean fundamental frequency and unusually rapid modulation in pitch—elicit greater skin conductance responses in adult listeners—an index of sympathetic arousal (Crowe & Zeskind, 1992). These observations of the coupled regulation of infant cardiac vagal regulation and acoustic vocal features of cries as well as their effects on the perceptions and physiological responses of listeners are consistent with the general mammalian principals of social communication outlined earlier.

These proposed general principles are also consistent with converging evidence from the study of the

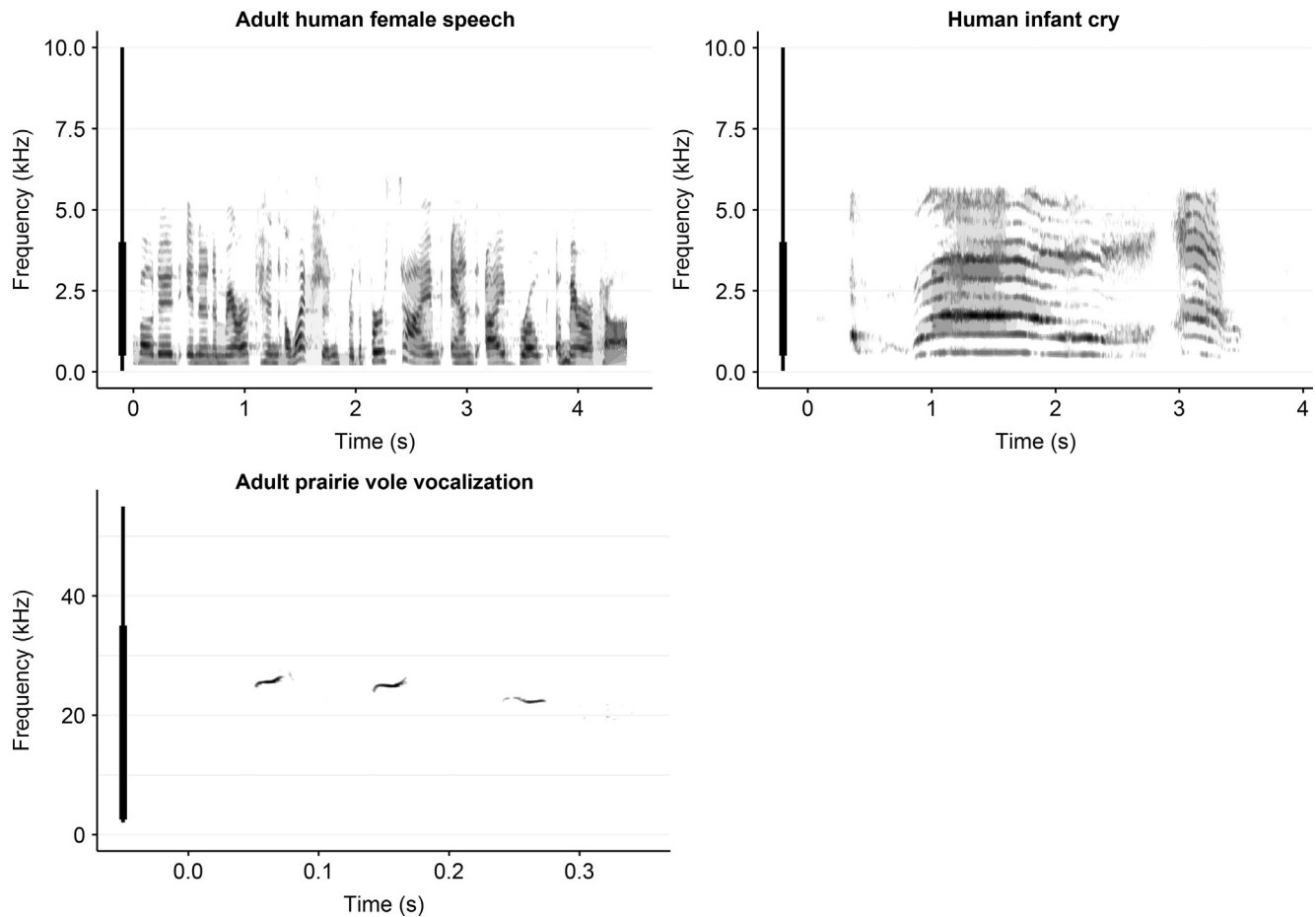


FIG. 3.3 Spectrograms of adult human female speech, human infant cry, and affiliative calls of an adult prairie vole (*Microtus ochrogaster*) with species-typical frequency band of perceptual advantage (thick bars).

vocalizations of African elephant (*L. africana*), who lead highly structured social lives with intergroup communication of affiliative and danger cues playing a key role in survival (Langbauer, 2000). As described earlier, information about middle ear resonance in elephants is lacking. While this limits the testing of specific hypotheses based on absolute values of the band of perceptual advantage, evidence from the African elephant can be illustrative for demonstrating the principles of vocal acoustic features and their parallel regulation with emotional states.

The most common elephant vocalization is the rumble (Berg, 1983; Leong, Ortolani, Burks, Mellen, & Savage, 2003; Stoeger-Horwath, Stoeger, Schwammer, & Kratochvil, 2007), a call with a typical fundamental frequency below 20 Hz (Soltis, 2010). Experimental induction of phonation using an excised elephant larynx is highly suggestive of elephant rumbles as produced by self-sustained vocal fold vibration caused by the exhalation of air from the lungs (Herbst et al., 2012), the same mechanism used to generate phonated speech in humans. As vocal fold tension and length directly

contribute to the fundamental frequency of a vocalization, it would be expected that elephant vocalization acoustic features would be regulated in concert with internal states via projections arising from the nucleus ambiguus. Although study of the neural regulation of emotional expression in elephants is currently scant, elephant behavior and situational cues may be used as a proxy for internal state regulation. Thus, we would predict that defense-inducing situations or displays of behavioral defensiveness would be related to vocalization fundamental frequencies and their modulation according to the same principles observed in infant cries. Indeed, adult female African elephants produce rumbles with higher and more variable fundamental frequencies when interacting with more dominant conspecifics—high arousal situations that are often coupled with visible markers of arousal such as urination, defecation, and fleeing—compared to rumbles in neutral social situations (Soltis, Leighty, Wesolek, & Savage, 2009).

In contrast to the very low fundamental frequencies of rumbles, African elephant roars are among the highest

pitched vocalizations these animals produce with fundamental frequencies above 300Hz and the first three resonating frequencies (formants) in the 450–1500Hz range (Stoeger, Charlton, Kratochvil, & Fitch, 2011). As would be expected from the relation between vocalizations and internal states, these relatively high-frequency roars are common under distress (Stoeger-Horwath et al., 2007). Infant African elephants in situations of behaviorally observed high arousal have higher levels of chaotic, nonlinear vocal production; these losses in tonality are associated with increased fundamental frequencies (Stoeger et al., 2011). This suggests that increased arousal in these calls may reflect unstable modulation of the vocalization structures, which may be induced by a decrease of neural regulation coordinated in part through the ventral vagal complex.

Further converging evidence for the principles of mammalian coregulation between vocalizations and biobehavioral states comes from the prairie vole (*M. ochrogaster*), a particularly useful rodent model for examining social biobehavioral state regulation. Behaviorally, prairie vole life is highly socially structured and includes monogamous pair bonding (Carter, Devries, & Getz, 1995). Their bodily states have strong influence via the myelinated vagus nerve, as evidenced by low heart rates for relative body size and high levels of heart rate variability in frequencies of spontaneous respiration (i.e., RSA; Grippo, Lamb, Carter, & Porges, 2007). Furthermore, vocalizations play a prominent role in their behavioral repertoire (e.g., Lepri, Theodorides, & Wysocki, 1988) and they feature a large auditory cortex

for vocal processing (Campi, Karlen, Bales, & Krubitzer, 2007). The prairie vole calls generated by the whistle mechanism described above have little to no harmonic power (Fig. 3.3, bottom panel). This observation suggests that social signals embedded in these calls may be largely determined by the regulation of the larynx and that the pharynx may play a small role, if any, in their modulation. Thus, fundamental frequency and its regulation by the branches of the vagus nerve may carry the majority of biologically significant information embedded in prairie vole vocalizations.

Fig. 3.4 displays the results of a reanalysis of data initially reported by Stewart et al. (2015) in which we examined state-related within individual covariation of RSA and vocalization fundamental frequency. In this experimental design, voles were briefly relocated from their home cage and then placed in a test cage that created a physical separation from a familiar vole but allowed the exchange of vocal cues. Mean RSA per minute—a measure of ventral vagal regulation of the heart—was extracted from implanted radiotelemetry transmitters that recorded continuous ECG signals (e.g., Grippo et al., 2007). Of the three voles included in the analysis, two had consistent and substantial negative relations between RSA and vocalization fundamental frequency (Fig. 3.4), with a 1-unit increase in RSA predicting an approximately 3500Hz decrease in mean vocalization fundamental frequency. Vocalization fundamental frequencies were near or above the upper bound of the prairie vole band of perceptual advantage (as estimated from the middle ear resonant frequency of *M. arvalis*) when

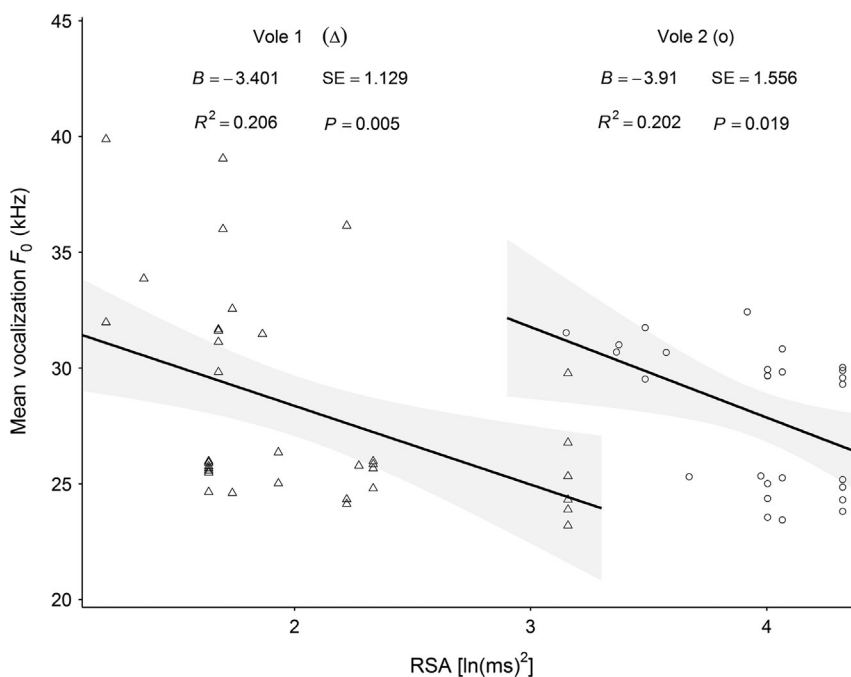


FIG. 3.4 Relation between vocalization fundamental frequency (F_0) and cardiac vagal tone, as indexed by respiratory sinus arrhythmia (RSA; method described in Porges, 1985), in two prairie voles (*Microtus ochrogaster*).

ventral vagal complex regulation of the heart was at its weakest (i.e., at the lowest RSA values).

V SUMMARY

The regulation of mean spectral and spectrotemporal modulation properties of vocalization as part of the ventral vagal complex provides a framework for interpreting patterns in the production and reception of mammalian vocal communication. We propose that the frequency band of perceptual advantage of individual mammal groups, constrained by the physical properties of middle ear structures, directly affects the frequency ranges in which cues of safety and danger are communicated between conspecifics. Rodents utilize a whistle mechanism for social communication within their frequency band of perceptual advantage as determined by the high resonant frequencies of their small middle ear bones. Larger mammals (e.g., humans) may rely on vocal fold vibration, a method that generates lower-frequency fundamentals more aligned with lower frequency bands of advantage.

The common neural ventral vagal system that links supradiaphragmatic organ regulation (heart and lungs) with muscles that produce and alter vocalization acoustic features (the larynx for whistles; the larynx and pharynx for vocalizations produced by vocal fold vibration) provides a mechanism for embedding information about internal states within frequency and spectrotemporal modulation features of vocalizations. In addition to the examples outlined in this chapter, emotional arousal coincides with increased fundamental frequencies in a range of mammalian groups (e.g., [Schehka & Zimmermann, 2009](#); [Zimmermann, 2009](#)). This embedded information facilitates shared state regulation

between conspecifics, allowing rapid proliferation of survival-related autonomic responses to danger cues as well as promotion of the affiliative behavior that is key to the socially dependent lives of mammals.

In our proposed model, listener reception of these vocal cues is integrated with neural efferent projections to the middle ear muscles, which dynamically regulate response sensitivity in the frequency band of perceptual advantage. Acoustic danger cues decrease sensitivity for social communication while increasing the sensitivity for low- and high-frequency sounds. Functionally, this promotes an increased sensitivity toward frequency bands that are associated with the relatively low-frequency calls of larger predators and the relatively high-frequency distress vocalizations of conspecifics. On the other hand, safety cues increase sensitivity for signals within the frequency band of perceptual advantage, downregulate the sensitivity to danger cues, and facilitate an efficient extraction of conspecific vocalizations from environmental noise ([Fig. 3.5](#)).

While our examples provide tentative support for this general framework of mammalian social communication of stereotyped calls, many aspects—such as the measurement of ventral vagal complex influences on human speech—remain untested. Furthermore, ecological pressures and evolutionary relations between prey and predators may further modify the frequency band of perceptual advantage independent of the middle ear resonant frequency (e.g., acoustic qualities of underground tunnel habits for subterranean dwellers, cats' heightened sensitivity for high-frequency sounds). Finally, the proposed neural mechanisms in this model are largely dependent on circuits in the brainstem and periphery. Further research is needed to examine the likely bidirectional interactions of these circuits with higher level cortical processing. Given the wide diversity of calls

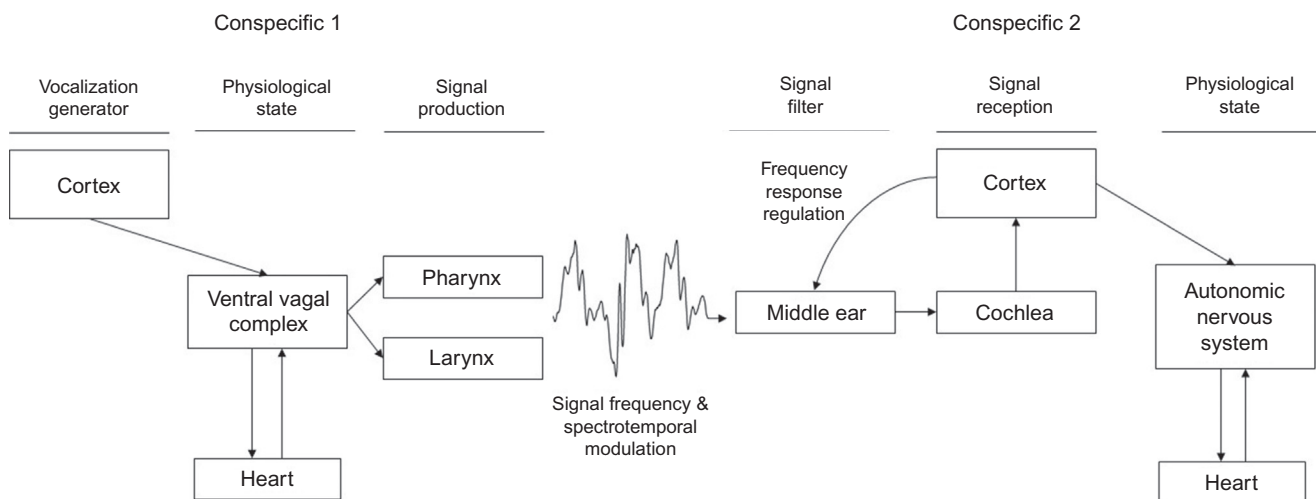


FIG. 3.5 Model of integrated vocal communication and physiological regulation between mammalian conspecifics.

observed across mammalian species, we do not expect this framework to describe all possible encoded and receptive features of vocalizations. However, we encourage researchers to apply this novel evolutionary biobehavioral framework to generate new research questions and hypotheses investigating the expression of physiological and emotional states in vocalizations and the influence vocalizations have on the physiological and emotional states of the listener.

References

- Berg, J. K. (1983). Vocalizations and associated behaviors of the African elephant (*Loxodonta africana*) in captivity. *Zeitschrift für Tierpsychologie*, 63(1), 63–79.
- Berke, G. S., & Long, J. L. (2010). Functions of the larynx and production of sounds. In S. M. Brudzynski (Ed.), *Handbook of mammalian vocalizations: An integrative neuroscience approach* (pp. 419–426). Amsterdam: Academic Press/Elsevier.
- Brown, M. C., & Levine, J. L. (2008). Dendrites of medial olivocochlear neurons in mouse. *Neuroscience*, 154(1), 147–159.
- Campi, K. L., Karlen, S. J., Bales, K. L., & Krubitzer, L. (2007). Organization of sensory neocortex in prairie voles (*Microtus ochrogaster*). *Journal of Comparative Neurology*, 502(3), 414–426.
- Carter, C. S., Devries, A. C., & Getz, L. L. (1995). Physiological substrates of mammalian monogamy: the prairie vole model. *Neuroscience and Biobehavioral Reviews*, 19(2), 303–314.
- Crowe, H. P., & Zeskind, P. S. (1992). Psychophysiological and perceptual responses to infant cries varying in pitch: comparison of adults with low and high scores on the Child Abuse Potential Inventory. *Child Abuse and Neglect*, 16(1), 19–29.
- Crumley, R. L. (1994). Unilateral recurrent laryngeal nerve paralysis. *Journal of Voice*, 8(1), 79–83.
- Fitch, W. T. (2006). Production of vocalizations in mammals. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed., pp. 115–121). Oxford: Elsevier.
- Grippe, A. J., Lamb, D. G., Carter, C. S., & Porges, S. W. (2007). Social isolation disrupts autonomic regulation of the heart and influences negative affective behaviors. *Biological Psychiatry*, 62(10), 1162–1170.
- Gustafson, G. E., & Green, J. A. (1989). On the importance of fundamental frequency and other acoustic features in cry perception and infant development. *Child Development*, 60(4), 772–780.
- Heffner, H. E., & Heffner, R. S. (2007). Hearing ranges of laboratory animals. *Journal of the American Association for Laboratory Science*, 46(1), 20–22.
- Heffner, R. S., & Heffner, H. E. (1982). Hearing in the elephant (*Elephas maximus*): absolute sensitivity, frequency discrimination, and sound localization. *Journal of Comparative and Physiological Psychology*, 96(6), 926–944.
- Herbst, C. T., Stoeger, A. S., Frey, R., Lohscheller, J., Titze, I. R., Gumpenberger, M., et al. (2012). How low can you go? Physical production mechanism of elephant infrasonic vocalizations. *Science*, 337(6094), 595–599.
- Hof, J. F. Y. (2010). Laryngeal muscles as highly specialized organs in airway protection, respiration and phonation. In S. M. Brudzynski (Ed.), *Handbook of mammalian vocalizations: An integrative neuroscience approach* (pp. 13–21). Amsterdam: Academic Press/Elsevier.
- Kolacz, J. (2016). Investigating infant crying persistence and cry acoustic features as early risk indicators for social adjustment: Developmental associations with infant vagal tone and attachment stress. (Doctoral dissertation).
- Kolacz, J., Lewis, G. F., Gariépy, J.-L., & Porges, S. W. (in preparation). Infant cry acoustic features: a window into the developing neural social engagement system.
- Kryter, K. D. (1962). Methods for the calculation and use of the articulation index. *The Journal of the Acoustical Society of America*, 34(11), 1689–1697.
- Langbauer, W. R. (2000). Elephant communication. *Zoo Biology*, 19, 425–445.
- Lange, S., Stalleicken, J., & Burda, H. (2004). Functional morphology of the ear in fossorial rodents, *Microtus arvalis* and *Arvicola terrestris*. *Journal of Morphology*, 262(3), 770–779.
- Leong, K. M., Ortolani, A., Burks, K. D., Mellen, J. D., & Savage, A. (2003). Quantifying acoustic and temporal characteristics of vocalizations for a group of captive African elephants *Loxodonta africana*. *Bioacoustics*, 13, 213–231.
- Lepri, J. J., Theodorides, M., & Wysocki, C. J. (1988). Ultrasonic vocalizations by adult prairie voles, *Microtus ochrogaster*. *Experientia*, 44(3), 271–273.
- Liberman, M. C., & Guinan, J. J., Jr. (1998). Feedback control of the auditory periphery: anti-masking effects of middle ear muscles vs. olivocochlear efferents. *Journal of Communication Disorders*, 31(6), 471–482.
- Lutman, M., & Martin, A. (1979). Development of an electroacoustic analogue model of the middle ear and acoustic reflex. *Journal of Sound and Vibration*, 64(1), 133–157.
- Markl, H., & Ehret, G. (1973). Auditory threshold of the mouse (*Mus musculus*). *Zeitschrift für Tierpsychologie*, 33, 274–286.
- Mukerji, S., Windsor, A. M. M., & Lee, D. J. (2010). Auditory brainstem circuits that mediate the middle ear muscle reflex. *Trends in Amplification*, 14(3), 170–191.
- Nummela, S. (1995). Scaling of the mammalian middle ear. *Hearing Research*, 85, 18–30.
- Orestes, M. I., & Chhetri, D. K. (2014). Superior laryngeal nerve injury: effects, clinical findings, prognosis, and management options. *Current Opinion in Otolaryngology and Head and Neck Surgery*, 22(6), 439–443.
- Pang, X. D., & Guinan, J. J. (1997). Effects of stapedius-muscle contractions on the masking of auditory-nerve responses. *The Journal of the Acoustical Society of America*, 102(6), 3576–3586.
- Patriquin, M. A., Scarpa, A., Friedman, B. H., & Porges, S. W. (2013). Respiratory sinus arrhythmia: a marker for positive social functioning and receptive language skills in children with autism spectrum disorders. *Developmental Psychobiology*, 55(2), 101–112.
- Porges, S. W. (1985). Method and apparatus for evaluating rhythmic oscillations in aperiodic physiological response systems. U.S. Patent 4510944.
- Porges, S. W. (1995). Orienting in a defensive world: mammalian modifications of our evolutionary heritage. A Polyvagal Theory. *Psychophysiology*, 32, 301–318.
- Porges, S. W. (2001). The polyvagal theory: phylogenetic substrates of a social nervous system. *International Journal of Psychophysiology*, 42, 123–146.
- Porges, S. W. (2004). Neuroception: a subconscious system for detecting threat and safety. *Zero to Three: Bulletin of the National Center for Clinical Infant Programs*, 24(5), 19–24.
- Porges, S. W. (2007). The polyvagal perspective. *Biological Psychology*, 74, 116–143.
- Porges, S. W. (2011). *The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication, and self-regulation*. New York: WW Norton.
- Porges, S. W. (2015). Making the world safe for our children: down-regulating defence and Up-regulating social engagement to 'optimise' the human experience. *Children Australia*, 40(02), 114–123.
- Porges, S. W., Bazhenova, O. F., Bal, E., Carlson, N., Sorokin, Y., Heilman, K. J., et al. (2014). Reducing auditory hypersensitivities

- in autistic spectrum disorder: preliminary findings evaluating the listening project protocol. *Frontiers in Pediatrics*, 2, 80.
- Porges, S. W., & Lewis, G. F. (2010). The polyvagal hypothesis: common mechanisms mediation autonomic regulation, vocalizations, and listening. In S. M. Brudzynski (Ed.), *Handbook of mammalian vocalizations: An integrative neuroscience approach* (pp. 255–264). Amsterdam: Academic Press/Elsevier.
- Porges, S. W., Macellaio, M., Stanfill, S. D., McCue, K., Lewis, G. F., Harden, E. R., et al. (2013). Respiratory sinus arrhythmia and auditory processing in autism: modifiable deficits of an integrated social engagement system? *International Journal of Psychophysiology*, 88, 261–270.
- Porter, F. L., Porges, S. W., & Marshall, R. E. (1988). Newborn pain cries and vagal tone: parallel changes in response to circumcision. *Child Development*, 59(2), 495–505.
- Rafal, Z. (2013). Conductive hearing loss in children with autism. *European Journal of Pediatrics*, 1(72), 1007–1010.
- Riede, T. (2013). Stereotypic laryngeal and respiratory motor patterns generate different call types in rat ultrasound vocalization. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 319(4), 213–224.
- Riede, T., York, A., Furst, S., Müller, R., & Seebeck, S. (2011). Elasticity and stress relaxation of a very small vocal fold. *Journal of Biomechanics*, 44(10), 1936–1940.
- Rouiller, E. M., Capt, M., Dolivo, M., & De Ribaupierre, F. (1989). Neuronal organization of the stapedius reflex pathways in the rat: a retrograde HRP and viral transneuronal tracing study. *Brain Research*, 476(1), 21–28.
- Roy, N., Smith, M. E., Dromey, C., Redd, J., Neff, S., & Grennan, D. (2009). Exploring the phonatory effects of external superior laryngeal nerve paralysis: an in vivo model. *The Laryngoscope*, 119(4), 816–826.
- Saigusa, H. (2010). Comparative anatomy of the larynx and related structures. *Japan Medical Association Journal*, 139, 803–808.
- Schekka, S., & Zimmermann, E. (2009). Acoustic features to arousal and identity in disturbance calls of tree shrews (*Tupaia belangeri*). *Behavioral Brain Research*, 203, 223–231.
- Shinya, Y., Kawai, M., Niwa, F., & Myowa-Yamakoshi, M. (2016). Associations between respiratory sinus arrhythmia and fundamental frequency of spontaneous crying in preterm and term infants at term-equivalent age. *Developmental Psychobiology*, 58(6), 724–733.
- Silman, S., & Silverman, C. A. (1991). *Auditory diagnosis*. San Diego: Academic Press.
- Singh, N. C., & Theunissen, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *The Journal of the Acoustical Society of America*, 114(6), 3394–3411.
- Soltis, J. (2010). Vocal communication in African elephants (*Loxodonta africana*). *Zoo Biology*, 29(2), 192–209.
- Soltis, P., Leighty, K. A., Wesolek, C. M., & Savage, A. (2009). The expression of affect in African elephants (*Loxodonta africana*) rumble vocalizations. *Journal of Comparative Psychology*, 123, 222–225.
- Stewart, A. M., Lewis, G. F., Heilman, K. J., Davila, M. I., Coleman, D. D., Aylward, S. A., et al. (2013). The covariation of acoustic features of infant cries and autonomic state. *Physiology and Behavior*, 120, 203–210.
- Stewart, A. M., Lewis, G. F., Yee, J. R., Kenkel, W. M., Davila, M. I., Carter, C. S., et al. (2015). Acoustic features of prairie vole (*Microtus ochrogaster*) ultrasonic vocalizations covary with heart rate. *Physiology and Behavior*, 138, 94–100.
- Stoeger, A. S., Charlton, B. D., Kratochvil, H., & Fitch, W. T. (2011). Vocal cues indicate level of arousal in infant African elephant roars. *The Journal of the Acoustical Society of America*, 130(3), 1700–1710.
- Stoeger-Horwath, A. S., Stoeger, S., Schwammer, H. M., & Kratochvil, H. (2007). Call repertoire of infant African elephants: first insights into the early vocal ontogeny. *The Journal of the Acoustical Society of America*, 121(6), 3922–3931.
- Zeskind, P., & Marshall, T. (1988). The relation between variations in pitch and maternal perceptions of infant crying. *Child Development*, 59(1), 193–196.
- Zimmermann, E. (2009). Vocal expression of emotion in a nocturnal prosimian primate group, mouse lemurs. In S. M. Brudzynski (Ed.), *Handbook of mammalian vocalizations: An integrative neuroscience approach* (pp. 215–225). Amsterdam: Academic Press/Elsevier.
- Zwislocki, J. J. (2002). Auditory system: peripheral nonlinearity and central additivity, as revealed in the human stapedius-muscle reflex. *Proceedings of the National Academy of Sciences of the United States of America*, 99(22), 14601–14606.