

COMPARATIVE COGNITION & BEHAVIOR REVIEWS

Elephants and Sirenians: A Comparative Review across Related Taxa in Regard to Learned Vocal Behavior

Anton Baotic

Department of Behavioral and Cognitive Biology, University of Vienna

Beth Brady

Mote Marine Laboratory

Eric Angel Ramos

Fundación Internacional para la Naturaleza y la Sustentabilidad, and The Rockefeller University

Angela S. Stoeger

Department of Behavioral and Cognitive Biology, University of Vienna

Vocal production learning is the ability to modify a vocal output in response to auditory experience. It is essential for human speech production and language acquisition. Vocal learning evolved independently several times in vertebrates, indicating evolutionary pressure in favor of this trait. This enables cross-species comparative analysis to be used to test evolutionary hypotheses. Humans share this ability with a versatile but limited group of species: songbirds, parrots and hummingbirds, bats, cetaceans, seals, and elephants. Although case studies demonstrate that African savanna and Asian elephants are capable of heterospecific imitation, including imitation of human words, our understanding of both the underlying mechanisms and the adaptive relevance within the elephant's natural communication system is limited.

Even though comparing phylogenetically distant species is intriguing, it is also worthwhile to investigate whether and to what extent learned vocal behavior is apparent in species phylogenetically close to an established vocal learner. For elephants, this entails determining whether their living relatives share their special ability for (complex) vocal learning.

In this review, we address vocal learning in Elephantidea and Sirenia, sister groups within the Paenungulata. So far, no research has been done on vocal learning in Sirenians. Because of their aquatic lifestyle, vocalization structure, and evolutionary relationship to elephants, we believe Sirenians are a particularly interesting group to study. This review covers the most important acoustic aspects related to vocal learning in elephants, manatees, and dugongs, as well as knowledge gaps that must be filled for one to fully comprehend why vocal learning evolved (or did not) in these distinctive but phylogenetically related taxa.

Keywords: elephants, manatees, dugongs, communication, vocal repertoire, vocal behavior, vocal learning

Introduction

Vocal learning is critical for the development of human speech and language. Although human language is unique, some animals share our capacity for vocal learning. Vocal learning is a behavioral trait in which an individual's vocalizations are shaped by their social and acoustic environment (Nottebohm, 1972).

We use the term *vocal learning* to refer to the various forms of learning based on the leading framework developed by Janik and Slater (1997, 2000). This framework categorizes *vocal learning* into vocal comprehension learning (VCL; Nowicki & Searcy, 2014), vocal usage learning (VUL), and vocal production learning (VPL). VCL occurs when an individual extracts a novel meaning from a signal. VUL appears when an individual learns to produce an existing signal in a novel context. For VPL to occur, a vocalization needs to be modified based on auditory experience.

In the literature, research has long suggested that VPL is restricted to species that can imitate novel sounds and have a direct neuronal connection between the forebrain and phonatory muscles as observed in the larynx in mammals or syrinx in birds (e.g., Fitch, 2010). Recent accumulating evidence from species exhibiting fine-tuning of acoustic features of the innate call repertoire, such as call convergence and social feedback contingency, started to challenge this assumption. As is now widely accepted, vocal learning in general should be viewed as a continuum, or as a multidimensional space with various complex levels that may also differ in the neural pathways involved (Jarvis, 2019; Tyack, 2020). For example, Vernes, Kriengwatana, et al. (2021), considered two dimensions for VUL: how vocalization context (behavioral and positional contexts) is learned and how timing (call timing and rhythmicity) is learned. For VPL, three dimensions were

defined: (a) the copying of auditory signals (considering copying accuracy but also divergence of acoustic models); (b) the degree and type of change or modification; and (c) timing with respect to when learning happens, how long it takes, and how long it is retained.

Other authors have recently put forward ideas and frameworks that go in similar directions, also suggesting that besides specialized neuronal implementation other criteria for learned vocal behavior and functional pressures need to be identified (e.g., Martins & Boeckx, 2020; Wirthlin et al., 2019), because neural circuits in mammalian species (e.g., dolphins, seals, elephants, manatees) are difficult to study. In many species, investigating potential vocal learning phenotypes based on, for example, behavioral ecology is a more promising strategy.

A comprehensive understanding of the complex vocal learning trait and its evolution requires a broad cross-species comparative approach focusing on specific research questions (Lattenkamp & Vernes, 2018). This comparative approach is critical because no single species is capable of providing all answers regarding behavioral expression, neuronal control, or functional relevance of vocal learning. For species that are recognized VPLers, to define their vocal learning phenotype in depth, three critical questions exist: *How* is VPL expressed? *What* do animals use it for within their natural communication system? *Who* else is capable of VPL phylogenetically close to the species in question (Lattenkamp, 2020)?

Species with the ability to copy novel sounds that are not part of the species-specific repertoire are considered complex VPLers (e.g., Petkov & Jarvis, 2012). Humans share the capacity for complex VPL with a diverse range of species. Oscine songbirds (Lipkind et al., 2020) and parrots (Pepperberg, 2010) can be trained to imitate human speech and have been shown to imitate vocalizations of other species or artificial sound models (Tyack, 2020). However, most species use VPL to acquire species-specific songs. Hummingbirds have been shown to replace their song when hearing a new song (Araya-Salas & Wright, 2013; Johnson & Clark, 2020). Cetaceans (Janik, 2014), pinnipeds (Reichmuth & Casey, 2014), and elephants (Stoeger & Manger, 2014) are capable of heterospecific vocal imitation, including imitation of human words and sounds resembling human speech in rhythm and frequency (Tyack, 2020).

Three different elephant species currently exist: the African forest elephant, *Loxodonta cyclotis*; African savanna elephant, *Loxodonta africana*; and the Asian elephant, *Elephas maximus*. Any knowledge or research on addressing VPL stems from African savanna and Asian

Author Note: Angela S. Stoeger, Department of Behavioral and Cognitive Biology, University of Vienna, Djerassiplatz 1, 1030 Vienna, Austria

Correspondence concerning this article should be addressed to Angela S. Stoeger at angela.stoeger-horwath@univie.ac.at

Acknowledgments: We thank the editor and two anonymous reviewers for their helpful suggestions and comments that improved the quality of the manuscript. We are grateful for Dr. Joy S. Reidenberg's intellectual input into the design of Figure 5, illustrating the manatee and elephant vocal apparatus.

Funding: Anton Baotic and Angela S. Stoeger are funded by the FWF, the Austrian Science Fund (P 31034).

elephants. For example, a male African savanna elephant named Calimero imitated the high-frequency squeaks typically produced by the Asian elephants (but not by wild African elephants) with which he grew up as the only African elephant. A female African savanna elephant imitated the sounds of trucks (Poole et al., 2005). A male Asian elephant named Koshik has been shown to imitate human speech, matching speech features in such detail that Korean native speakers could readily understand and transcribe the imitations (Stoeger, Mietchen, et al., 2012). The factors that may have led Koshik to imitate his human caregivers may be social deprivation from conspecifics during an important period of bonding and development when humans were the only social contact available. This hypothesis may also hold true for Calimero, and for other known examples of speech imitation in mammals such as Hoover the seal (Ralls et al., 1985), the beluga named Logosi (Eaton, 1974), and most talking birds (e.g., parrots; Pepperberg, 2010).

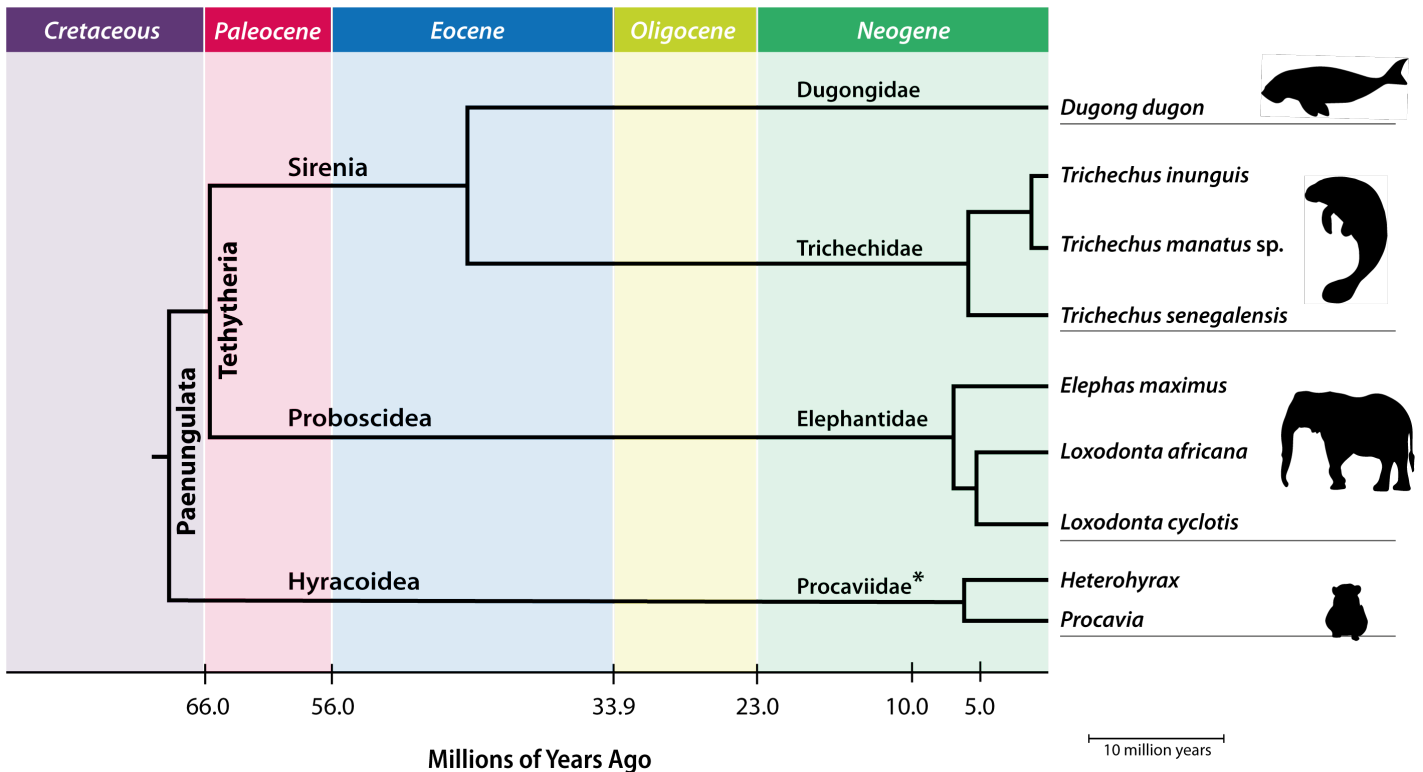
VUL is more widespread (Shapiro et al., 2004) and has been demonstrated in multiple species, including parrots, chickens, mynahs, cats, dogs, rats (Vernes, Kriegswatana, et al., 2021), monkeys and apes (Cheney & Seyfarth, 2018), whales and dolphins (Janik, 2014), pinnipeds

(Reichmuth & Casey, 2014; Shapiro et al., 2004), and bats (Lattenkamp et al., 2018). VUL can experimentally be demonstrated if an animal is able to vocalize in response to a conditioning stimulus (Janik & Slater, 2000; Shapiro et al., 2004). Stoeger and Baotic (2021) and Stoeger et al. (2021) showed that adult African savanna elephants can vocalize in response to verbal cues, reliably producing call types normally used in social communication as well as idiosyncratic sounds. The latter are produced via unusual and individualistic production techniques applying non-phonatory structures (e.g., nasal tissue vibration via an ingressive airflow at the trunk tip) or by contracting defined superficial muscles at the trunk base (Stoeger & Baotic, 2021; Stoeger et al., 2021). This reveals a profound vocal control over diverse sound production mechanisms.

Twenty years after the discovery that elephants are vocal learners, our understanding of the vocal learning phenotype remains limited. We know nothing about the adaptive function(s) of the vocal learning ability within the natural communication system and have no knowledge about the vocal learning ability, or lack thereof, of the elephants' closest relatives.

Elephants (*Elephantidae*, *Proboscidea*), belong to the *Afrotheria*, a superorder that represents a diverse group

Figure 1. Phylogenetic tree of Paenungulata including time divergence between Sirenia, Proboscidea, and Hyracoidea. Divergence dates at nodes are in millions of years, based on Springer et al. (2015), Meyer et al. (2017) and de Souza et al. (2021). Note that Procaviidae* are not discussed further in this review.



Vocal Production Learning (VPL)	YES	YES	???	???
Vocal innovation	yes	???	???	???
Vocal imitation	yes	yes	???	???
Contextual Learning	YES	???	???	???
Vocal Comprehension Learning (VCL)	???	???	???	???
Vocal Usage Learning (VUL)	yes	???	???	???
Functional relevance	???	???	???	???



African savanna elephant



Asian elephant



West Indian manatee



Rock hyrax

Figure 2. Knowledge about vocal learning in selected species of Paenungulata. The closest extant relatives of elephants are the aquatic dugongs and manatees and the terrestrial hyraxes (Kellogg et al., 2007). Very few studies have been conducted on Paenungulata vocal learning. Therefore, strong evidence is mostly missing, although observations of their vocal behavior indicate that learning might shape vocal communication in these taxa.

of mammals composed of six orders. Besides elephants, *Afrotheria* comprises sea cows (*Sirenia*), hyraxes (*Hyracoidea*), the armadillo (*Tubulidentata*), elephant shrews (*Macroscelidea*), golden moles and tenrecs (*Afrosoricida*) (van Dijk et al., 2001). Within the *Afrotheria*, *Proboscidea*, *Sirenia*, and *Hyracoidea* constitute the *Paenungulata*. However, branching within the *Paenungulata* is still debated (a phylogenetic time tree is provided in Figure 1). Rapid radiation, a deep divergence, and an extensive morphological diversification has led to limited phylogenetic signal confounding resolution at morphological and nucleotide levels (Amrine-Madsen et al., 2003; Kellogg et al., 2007; Nishihara et al., 2005; Pardini et al., 2007; Seiffert, 2007). Based on nine molecular loci, *Sirenia* is sister to *Hyracoidea* plus *Proboscidea* (Pardini et al., 2007), whereas morphological evidence supports a *Sirenian-Proboscidean* clade (*Tethytheria*) within *Paenungulata* (Tassy & Shoshani, 1988). *Tethytheria* (*Proboscidea* and *Sirenia*) is accepted by Asher et al. (2003), Kjer and Honeycutt (2007), and Seiffert (2007). Other researchers reject *Tethytheria* in favor of a *Hyracoidea-Sirenia* clade (Matthee et al., 2007; Springer & Murphy, 2007) or reject *Tethytheria* because of a suggested *Hyracoidea-Proboscidea* clade (Amrine-Madsen et al., 2003).

Sirenia is a clade of fully aquatic mammals comprising the three extant species of manatee (West Indian manatee, *Trichechus manatus*; African manatee, *T. senegalensis*; and Amazonian manatee, *T. inunguis*) and the dugong (*Dugong dugon*). In all, except the African manatees, acoustic

communication has been shown to occur in contexts of social and nonsocial aggregations and maternal care (Anderson & Barclay, 1995; O'Shea & Poché, 2006; Sousa-Lima et al., 2002). Although next to nothing is known about their vocal flexibility, studies carefully suggest that ambient noise levels might have an effect on manatee and dugong call usage and that both species modify their vocalizations to some extent (e.g., call rate, duration, call type usage) as a function of noise (Ando-Mizobata et al., 2011; Miksis-Olds & Tyack, 2009). In Florida manatees, a subspecies of the West Indian manatee, Miksis-Olds and Tyack (2009) showed that call types decreased during feeding and social behaviors in conditions of elevated noise levels (i.e., during nonwinter months and turbid environments). They found a decrease in vocalization rate dependent on behavioral state, an increase in chirp duration in the presence of calves, and an increase of squeak duration as noise level increased when calves were absent. This indicates not only that ambient noise levels have an effect on Florida manatee communication but also that Florida manatees are able to modify their vocalizations in response to external stimuli. The findings cautiously suggest their flexibility in call usage and structure, as a function of behavior and group composition.

Hyracoidea (rock hyrax and tree hyrax) are highly vocal, group living mammals. Adult males produce elaborate vocalizations known as songs. Songs are used as advertisement signals and possess geographical dialects with regard to syntax and syllable order (Kershenbaum et al., 2012). Research suggests that dispersing males introduce

song features to new areas, which are then learned (copied), repeated, and modified by locals, leading to individual and within-regional variation in song structure. This variation in song structure could represent a complex form of vocal learning. Besides these observations, VPL as well as VUL has not been specifically tested, and strong evidence in both these taxa is missing. Thus far, *Sirenia* and *Hyracoidea* are not considered among the species capable of vocal learning (Figure 2).

Although *Hyracoidea* is an important and relevant species within the *Paenungulata*, hyrax vocal signals (i.e., the elaborate songs with syllable and syntactic structure) are considerably different from signals that elephants, manatees, or dugongs produce. Therefore, in this article we review and compare what is known about elephant and Sirenian vocalizations. Specifically, we address vocal communicative behavior in the context of their socio-behavioral ecology, including sound production; the vocal repertoire and the information conveyed via vocalizations; and the vocal ontogeny and mother–infant behavior. The article’s intention is not to present a complete review of elephant and manatee vocal behavior, respectively. We aim to provide a synopsis of, in our opinion, the most relevant topics for future comparative research and to highlight knowledge gaps that need to be addressed in order for individuals to understand why certain vocal abilities—and, most important, vocal learning—did (or did not) evolve in these distinct (considering in terms of morphology and habitat use) but phylogenetically related taxa.

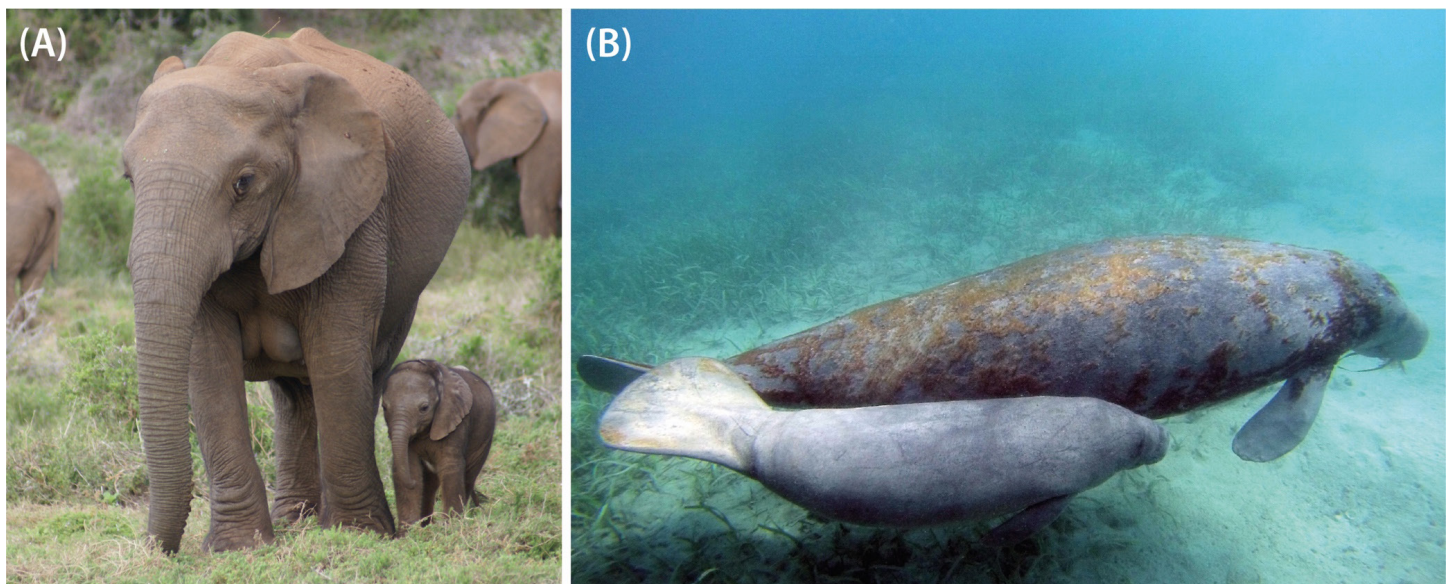
Social and Behavioral Ecology

Vocal communication is a form of social interaction, influenced by the social system and the social environment individuals live in. The social system and its complexity might be a driver of communication (and a possible causal factor in origins of human language) because of the need to mediate interactions and relations among group members (Freeberg et al., 2012). Thus, knowledge about animal social systems is an important prerequisite for the study of vocal communication.

African (*L. africana* and *L. cyclotis*) and Asian (*E. maximus*) elephants are generalist herbivores, long-lived, and highly social mammals. Females and males are sexually dimorphic in size (e.g., males are taller and bulkier). The females maintain extensive networks of social relationships between related matrilineal groups (McComb et al., 2000). These matriarchal family herds live in dynamic fission–fusion societies (Moss & Poole, 1983; Vidya & Sukumar, 2005), which are characterized by temporary merging of and splitting into family units that consist of groups of approximately 20 individuals (*L. cyclotis*, Turkalo et al., 2013; Turkalo & Fay, 1995; *L. africana*, Archie et al., 2006; Moss & Lee, 2011; Wittemyer et al., 2005).

Elephants show an extended period of maternal care. Born into stable family units, elephant calves are highly dependent upon their mothers for milk for approximately 2 years (Lee & Moss, 2011). Although the mother–calf bond in elephants is very strong (Figure 3A), calves are also frequently near other family members, such as sisters,

Figure 3. Photographs illustrating the strong mother–calf units in (A) African savanna elephants taken at the Addo Elephant National Park in South Africa (credit: Simon Stoeger), and (B) Antillean manatees, taken at Turneffe Atoll in the offshore of Belize in Central America (credit: Alton Jeffords).



aunts, great-aunts, or grandmothers; these members occasionally nurse closely related calves (*L. africana*, Lee, 1987; *E. maximus*, Vidya, 2014). On occasion, other adult females and juveniles guard or instruct calves, reflecting the highly social and cohesive nature of elephant families (Lee & Moss, 2011). Unlike the social-group-living females, pubertal males disperse from their natal family at an average age of 14 (Lee et al., 2011; Poole, 1989). African and Asian elephants exhibit a polygynous mating system, in which adolescent males are known to associate with other bulls, depending on their age and sexual state, and may rove between female groups throughout the year to temporarily associate and start competing with one another over receptive females (Eisenberg et al., 1971; Moss & Poole, 1983; Poole, 1989).

In Sirenians, complex social structures have not been identified yet. Although manatees have been described as semisocial (Hartman, 1979), manatees are more frequently observed in groups than alone (O'Shea & Poché, 2006; Reynolds, 1981). Large aggregations of Florida manatees occur during winter months at warm water refuges because of their inability to thermoregulate (Reynolds & Wilcox, 1994). Short-term, smaller aggregations have been observed during summer months in areas where they can find their primary resources: subaquatic vegetation and fresh water. Although these aggregations are triggered by nonsocial resources, they may provide an opportunity to socially engage with other individuals. However, these groups tend to be maintained for short durations, from hours to a few days (Reynolds, 1981). Lack of selective pressures may also have shaped the manatee's simple social structure (O'Shea & Poché, 2006; Reynolds, 1981). Manatees and dugongs are primarily herbivores, which does not necessitate chasing prey or cooperative foraging. Although reports of predation by sharks and alligators exist (Mou Sue et al., 1990; Wells et al., 1999; Wirsing et al., 2007a, 2007b), it is thought that dugongs and manatees minimize this risk because of their large size, sensory capabilities, thick skin, and behavioral changes when predators are perceived (Marsh et al., 2011). Manatees largely avoid predation, whereas dugongs are regular prey for various species of shark in the Indo-Pacific Ocean (Wirsing et al., 2007b, 2007a). Similar to elephants, the strongest social bond in manatees is the mother–calf bond (Figure 3B); the manatee has a long period of dependency at 1½ to 2 years (O'Shea & Poché, 2006; Reynolds, 1981). Manatee calves are thought to be nutritionally independent by 1 year of age but stay with their mother beyond this age, possibly to learn the location of warm water refuges and food sources (Deutsch et al., 2003).

Mating strategies vary slightly between manatees and dugongs. Both manatees and dugongs adapted scramble promiscuity (Anderson, 2002; Bengtson, 1981; Hartman, 1979), as male manatees have the opportunity to mate with multiple females but have little assurance of paternity. Reproductive behavior consists of mating herds, where a focal female is pursued by multiple males. In addition to mating herds, there have been observations of lekking behavior in which dugongs continuously occupied individual sites that they patrolled regularly and vocalized frequently during patrols (Anderson, 1997; Anderson & Barclay, 1995).

Structure and Function of Vocalizations

Identifying VUL or VPL in a particular species is challenging. One of the main problems is determining what represents a novel sound pattern, a novel call structure, or a novel sound in general. Thus, determining whether an observed vocalization is novel or represents a modification of an existing vocalization involves knowing the vocal repertoire of the species (Vernes, Kriengwatana, et al., 2021), the vocal range (in terms of call duration and frequency) and (if investigating VUL) the functionality of vocalizations. In elephants, we have considerable knowledge of the basic vocal repertoire, its variability, and the functional use of most vocalization types. In manatees and dugongs, the knowledge is currently lacking, and many aspects of the behavioral ecology are yet to be investigated.

Elephant vocalizations are distinguishable based on their frequency range (high or low frequency) and presence or absence of harmonics. Elephants commonly produce a set of structurally different calls. These include low-frequency *rumbles* with a fundamental frequency (f_0 , lowest frequency of a periodic sound) in the infrasonic range; midfrequency ranged *snorts*, *roars*, and *barks*; and (for elephants) high-frequency *trumpets* and *squeaks* (Stoeger & de Silva, 2014; spectrograms of representative call types are provided in Figure 4). These broad call types are produced in many different contextual situations (Poole, 2011). The *trumpet*, however, is the most characteristic utterance associated with elephants. All three species produce this high-frequency (f_0 range = ~300–800 Hz), loud, and bugling call type, particularly in situations of excitement or distress (e.g., greeting ceremonies between herds or births) but also during defensive or agonistic interactions (e.g., predator defense, mock charging; Leong et al., 2003; Nair et al., 2009; Poole, 2011). Usage and functionality may differ, however, as African elephants appear to trumpet and Asian elephants tend to roar more often in situations of arousal (Stoeger & de Silva, 2014).

Another high-pitched call type is the *squeak* (f_0 mean = 813.07 Hz), which is produced by Asian elephants via lip buzzing (Beeck et al., 2021; Stoeger & de Silva, 2014) in alarming or socially arousing contexts (de Silva, 2010; Herler & Stoeger, 2012; Nair et al., 2009). Squeaks have been heard from a few African savannah elephants; however, these are either imitations (of Asian elephant squeaks) or sound creations, respectively (Stoeger et al., 2021). The squeaks are made by squeezing the trunk tip together and closing one nostril while sucking in air through the other during inhalation. Both high-frequency trumpets and squeaks are periodic vocalizations with a coarse harmonic content. Roars and barks tend to occur in contexts of aggression or distress, such as during mating or separation between group members (Leong et al., 2003; Nair et al., 2009; Poole, 2011; Stoeger & de Silva, 2014). Roars often possess a rather chaotic acoustic structure, also referred to as *deterministic chaos* (Fitch et al., 2002). Barks in African and Asian elephants are not always chaotic. Although barks in Asian elephants are described as chaotic broadband vocalizations (with no fully clear harmonic structure), barks in African elephants sometimes have harmonic content (e.g., Stoeger et al., 2021).

Vocalizations in mammals that have noisy components (i.e., irregularities in the vibration of vocal folds) and are produced during different high-stakes contexts may be indicative of aroused physiological states (e.g., sexual arousal, threat arousal), which in turn may influence phonatory events. Particularly, nonlinear phenomena are considered to be honest signals that can serve as useful indicators for a caller's level of arousal (for a review, see Fitch et al., 2002). For instance, although noisy elephant roars are distress calls used in the context of separation from group members or in need of immediate care, tonal and mixed roars in elephant calves are produced in suckling contexts as protest calls when disturbed (Stoeger-Horwath et al., 2007).

Within each species, elephants are known to combine or concatenate *roars* and low-frequency *rumbles* in a variety of orders to produce “combination calls.” African savanna elephants and Asian elephants mostly emit roar–rumble combinations; *African forest elephants* produce more combinations (rumble–roars, rumble–roar–rumbles, roar–rumbles; Pardo et al., 2019). Although further behavioral studies are needed, Pardo et al. (2019) cautiously suggested based on their observations that call combinations in adults with different orders (particularly in the forest elephants) may have different functions (that differ across species). Although elephants produce a wide range of different vocalizations, the most common vocalization is the rumble.

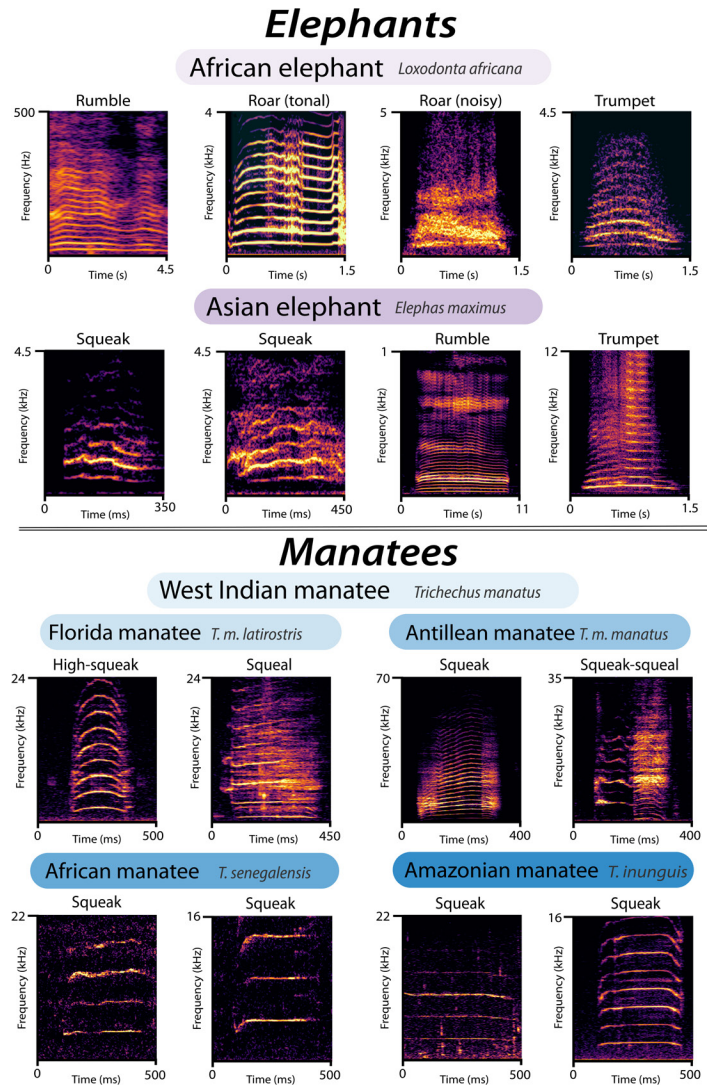


Figure 4. Comparative spectrograms of the different vocalizations in vocal repertoire of elephants and Sirenians. Examples of different call types produced by each species are displayed to illustrate the common structure of these calls and the frequency bands they occupy.

Rumbles are low-frequency and harmonically rich calls, with frequency components near or in the infrasonic range (Poole, 2011; Poole et al., 1988) and with variants that fall on a continuum (Soltis, 2010). Rumble vocalizations are uttered in multiple contexts, which include short- and long-distance communication. They range from coordinating movement and group activities, maintaining family group cohesion, keeping vocal contact with affiliates, and prompting defensive and exploratory behavior (Leighty et al., 2008; McComb et al., 2000, 2003; Poole, 2011).

The vocal repertoire of Sirenians varies between three and six call types depending on the mode of classification (Anderson & Barclay, 1995; Brady et al., 2020; Steel, 1982; Umeed et al., 2018). However, for most species, we

lack sufficient data to adequately characterize the breadth and function of their vocal repertoire. Most of the information about Sirenian vocal behavior comes from studies of the two West Indian manatee subspecies: the Florida and Antillean manatees. Brady et al. (2020) described their vocal repertoire as graded, with calls ranging from tonal sounds with harmonics that can extend into the ultrasonic range (> 20 kHz; Ramos et al., 2020) to atonal sounds that lack harmonic components (Mann et al., 2006).

The squeak is the most frequently described call type in manatees. Studies suggest that all Sirenian species make squeaks (referred to as *chirp-squeaks* in dugongs) that vary slightly in their range of f_0 (Anderson & Barclay, 1995; Nowacek et al., 2003; Rycyk et al., 2021; Sousa-Lima et al., 2008). West Indian manatees, African manatees, and dugongs produce vocalizations with overlapping f_0 ranges from 0.5 to 5 kHz (Anderson & Barclay, 1995; Nowacek et al., 2003; Rycyk et al., 2021), whereas Amazonian manatees produce calls with a higher f_0 at 1.6 to 8 kHz (Sousa-Lima et al., 2008). The other common vocalization described in the dugong and manatees is the atonal vocalization. Atonal vocalizations are referred to as *barks* in dugongs (Anderson & Barclay, 1995) or *squeals* in West Indian manatees (Brady et al., 2020). These broadband calls range in frequency from 0.5 Hz to 18 kHz (Anderson & Barclay, 1995; Miksis-Olds & Tyack, 2009) and possess deterministic chaos throughout the duration of the call (Mann et al., 2006). Spectrograms of representative calls are provided in Figure 4.

Although multiple call categories have been defined, there is limited information on the behavioral context of Sirenian vocalizations. Anderson and Barclay (1995) reported that specific call types were associated with territoriality, affiliative functions, and aggression in dugongs. In manatees, vocalizations have been noted during activities such as resting and feeding and when animals were engaged in cavorting (play) and mating (Bengtson & Fitzgerald, 1985; O'Shea & Poché, 2006). All ages and sexes have been observed in cavorting groups; however, most of these groups contain males (Marsh et al., 2011). The use of their vocalizations is thought to be affiliative and function primarily to maintain contact between mother and calf (Sousa-Lima et al., 2002). Spectrograms of manatee vocalizations show variability in call parameters; hypotheses state that these properties are related to changes in motivational state (O'Shea & Poché, 2006). For example, slight increases in duration of vocalizations from a mother–calf pair was noted when their behavior changed from resting to fleeing from a disturbance (O'Shea & Poché, 2006).

Some vocalizations of the repertoire from elephants (i.e., the trumpets and squeaks) as well as most manatee

and dugong sounds are considerably higher in frequency than expected for animals of that body size. Based on data resulting from phylogenetic regression, Ravignani and Garcia (2021) suggested a potential link between VPL and allometric outliers. The authors found that multiple species of VPL clades deviated from allometric scaling and show higher vocalization frequencies than expected for body size. Indeed, Ravignani and Garcia proposed an acoustic allometry hypothesis and suggested that *Trichechus inunguis* and *Trichechus manatus* may be potential candidates for VPL based on the high-frequency content in their squeak vocalizations.

Sound Production

The ability to vocally modify existing sounds and match a sound model is not only constrained by neuronal circuitry but also affected by the capabilities of the vocal production apparatus (Vernes, Kriengwatana, et al., 2021). It is crucial to consider and understand sound-producing mechanisms and peripheral morphological (including synapomorphic) structures of the species in question to assess constraints or, in some cases, a special ability for producing and modifying sounds (e.g., possessing a highly flexibly proboscis).

Mammal vocal production follows the source-filter theory, in which the vocal fold vibrations in the larynx (Fitch, 2006b), or vibration of different morphological structures (e.g., in dolphins; Madsen et al., 2012), generate a source signal determining the fundamental frequency. Usually, this signal is filtered by a vocal tract the shape and length of which amplifies certain resonant frequencies (spectra peaks or “formants”) before the vocalization radiates into the environment (Fitch, 2000).

Elephants hold a unique position among mammals because of their body size and the proboscis, a muscular hydrostat that elephants use extensively for sound emission and production. However, most knowledge on sound production stems from the rumbles, which have been shown to originate via passive vocal fold vibration (Herbst et al., 2012) and can be emitted orally as well as nasally (Stoeger, Heilmann, et al., 2012). Although the mode of rumble production is understood, the mechanism and anatomical sound-producing structures behind the higher-frequency vocalizations remain speculative, because the elephant's massive vocal folds can most likely not generate these high frequencies. A recent study suggested that Asian elephants produce their species-specific high-pitched squeaks (with a f_0 of up to 2000 Hz) by self-sustained lip vibrations when forcing air through tensed lips (also termed *lip buzzing*;

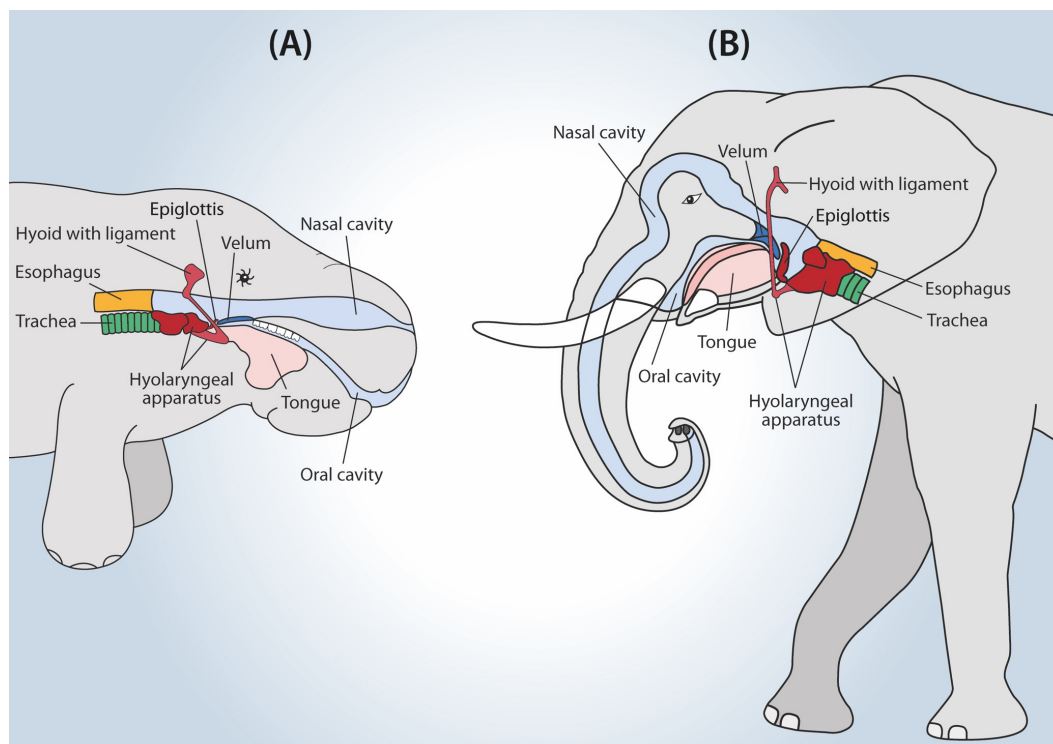
Beeck et al., 2021). In addition, African savanna elephants produce a variety of idiosyncratic sounds (Poole, 2011; Stoeger & Manger, 2014). Some individuals have been observed applying nasal tissue vibration via an ingressive airflow at the trunk tip to produce extraordinary high-frequency sounds (reaching a f_0 of up to 1800 Hz; Stoeger et al., 2021). These examples demonstrate the elephant's ability to use alternative sound production strategies to overcome morpho-mechanical limitations of laryngeal sound production to extend the frequency range available for communication (Beeck et al., 2021; Stoeger & Baotic, 2021; Stoeger et al., 2021) with a mechanism that potentially combines motoric abilities with vocal learning processes.

Vocal production in Sirenians is presumed to originate from the vocal folds in the laryngeal region (Grossman et al., 2014; Landrau-Giovanetti et al., 2014). The vocal folds are thick and are thought to regulate airflow to generate sound (Landrau-Giovanetti et al., 2014). As sound passes the source, it is thought that the sound resonates in the nasal region, as evidenced by observations of deflation and inflation of the nasal area while vocalizing (Grossman et al., 2014). Similarly, when elephants rumble intensively via the nasal vocal tract, a fluttering can be observed at

the elephant's forehead where the nasal passages enter the skull, which has also been used to identify vocalizing individuals (Payne et al., 1986). Vocalizations produced via the nostrils in Sirenians are unlikely, as there have been no structures capable of vibrating noted in the nasal cavity (Landrau-Giovanetti et al., 2014). The direct pathway from the source into the environment in Sirenians is not clear but is thought to be transmitted through either the soft tissues (Chapla et al., 2007) or the lingual and cervical fatty tissue of the head (Landrau-Giovanetti et al., 2014). Despite the limited knowledge of sound production in *Sirenia* and the anatomical differences, there are similarities in *Elephantidea* and *Sirenia* (Figure 5): highly flexible and multifunctional orofacial anatomical structures that have potential to alter the nasal and oral vocal tracts to modify the vocal output, or to generate sounds.

This anatomical structure has also been observed in walrus that possess tremendous vocal plasticity. Besides the larynx, walrus use their pharyngeal pouches, teeth, nose and mouth, the lips, and the mobile tongue to produce sounds (Schusterman & Reichmuth, 2008). Elephants also make use of their lips (Beeck et al., 2021), tongue, and the multifunctional trunk (Stoeger et al., 2021) to modify

Figure 5. Schematic main components of the (A) manatee vocal apparatus and (B) elephant vocal apparatus. In some sounds, the source signal is generated at the larynx and then filtered in the nasal and oral cavities (separated by the velum) of the vocal tract before radiating into the environment. Figure 5A styled from Figure 5 in Rommel et al. (2018). Figure 5B adapted from Figure 2A in Herbst et al. (2012) and from "A Novel Theory of Asian Elephant High-Frequency Squeak Production," by V. C. Beeck, G. Heilmann, M. Kerscher, and A. S. Stoeger, 2021, BMC Biology, 19, Article 12, Figure 1 (<https://doi.org/10.1186/s12915-021-01026-z>). Copyright 2021 by Springer Nature.



and to generate sounds. Manatees possess a highly flexible mouth with prehensile lips; the upper lip pad is split into left and right sides, which can move independently (Noll, 1983). Although they use it predominantly for browsing, these flexible structures could, in theory, be used for sound production. So far, we do not know whether manatees, similar to elephants, use synapomorphic structures to produce or modify sounds.

The production of complex vocalization in humans, birds, and nonhuman mammals always employs a sound source and filtering systems that need to be modified to match an acoustic template (orofacial structures, the tongue, nasal passages, jaws, cheeks or beaks; Vernes, Kriengwatana, et al., 2021). When addressing vocal learning in a cross-species comparative and phylogenetic approach, our interest cannot be limited to vocalizations that require control of the phonatory muscles. In mammals, the larynx is not necessarily the primary sound-producing structure (e.g., it is not the case in dolphins and some elephant vocalizations). Therefore, it is important and justified to focus research on vocalizations generated by the vocal organ and on supralaryngeal vocal production. Yet the question remains whether and how neural control mechanisms of VPL in laryngeal and nonlaryngeal sound production compares (Vernes, Kriengwatana, et al., 2021).

Vocal Ontogeny

Vocal ontogeny is another crucial aspect to consider when addressing vocal learning, yet the question of how much learning affects vocal development is difficult to resolve. During development and various stages of life, animals might change calls they use, where VUL may or may not be involved. Maturation processes lead to changes in vocal tract morphology that can naturally influence the structure of vocalizations. Therefore, observed changes in the vocal structure, or even in the vocal repertoire during ontogeny, are difficult to interpret in purely observational studies (Janik & Slater, 1997). During an individual's life, call structure might further change because of hormonal or other internal processes without learning being involved. Still, vocal ontogeny needs to be intensively studied in potential vocal learning species, as young individuals might be particularly flexible, and vocal learning might occur during limited sensitive phases (Vernes, Kriengwatana, et al., 2021).

Unfortunately, little is known about the vocal ontogeny of elephants across all species. Elephant calves generally have the same call type categories as adults, but their acoustic structure is different. Neonates roar and rumble

(though calf rumbles differ from adult rumbles, as they are higher in frequency, shorter in duration, and limited in modulation) soon after birth (Poole, 2011; Stoeger-Horwath et al., 2007). Proper trumpets appear to be created only after a few weeks of age, mostly in the context of playing. Before that, calves can blast air out of the trunk during play, but it sounds more like a snort than a trumpet (Stoeger-Horwath et al., 2007). In general, young elephants have a higher f_0 and a shorter call duration, which is true for the majority of call types, including rumbles and trumpets. Age-dependent variation (calculated in months) was documented in the low-frequency rumble. Older individuals (age 10 months and older) tended to have lower fundamental frequencies in rumbles than younger individuals (6 months and younger), and rumble duration tended to increase slightly with age. This finding is mainly explained by the smaller body size and related sound-producing structures as well as smaller respiratory capacities of young elephant calves. Infant African savanna elephants younger than 2 to 3 months of age can be heard producing soft grunts that are barely audible for observers but might function in close-contact communication between a mother and her calf. These quiet vocalizations might be a form of acoustic crypsis; recently, acoustic crypsis was suggested to occur between mother–calf pairs in whales (e.g., humpback whales, Videsen et al., 2017; southern right whale, Nielsen et al., 2019; North Atlantic right whales, Parks et al., 2019). Baleen whale mother–calf units vocalize at a particularly low level (resulting in a small active space; less than 100 m in humpback whales; Videsen et al., 2017) to minimize the risk of alerting potential predators such as eavesdropping killer whales. At approximately 2 months of age, elephant calves stop grunting (Stoeger-Horwath et al., 2007). Roars, barks, and rumbles, on the other hand, are produced frequently as they mature. Further, African savanna elephant calves and juveniles are more likely than adults to produce combination calls with varying call order (rumble–roar or roar–rumble–roar, and bark–rumble; Pardo et al., 2019; Stoeger et al., 2011). Adult African savanna elephants mostly produce roar–rumble combinations, which might cautiously hint toward a plastic stage with increased vocal variability during ontogeny.

Calves and juveniles of Asian elephants rumble, roar, trumpet, squeak, and produce variable combination of rumbles and roars as well (Herler & Stoeger, 2012; Pardo et al., 2019). The squeak's f_0 needs to be changed to a subscripted "0": f_0 does not appear to obey the inverse frequency–body mass rule. Adult males occasionally squeak at a higher frequency than calves, which may be explained by the call type's unique lip buzzing mode of sound production (Beck et al., 2021).

The vocal ontogeny of manatees is poorly understood. Manatee calves produce vocalizations that are structurally similar to those of adults. In both Antillean and Amazonian manatees, an inverse relationship between total body length and f_0 range was observed, indicating that the f_0 becomes more defined and lower as the animal ages (Sousa-Lima et al., 2002, 2008). Additionally, manatee calves may exhibit structural features that are recognizable by unrelated conspecifics. Playback experiments of a distressed calf call to unrelated females appeared to elicit a protective response, where females encircled the source of the sound with their tails facing the speaker (Phillips et al., 2003). Although the acoustic parameters of the call were not reported, it is possible that the call type played had features those females recognized as a calf.

It is crucial to understand why changes occur, retaining a critical and careful perspective, because babbling could also be an explanation for changes in utterances. Babbling is a developmental phenomenon known in human vocal learning and found in many nonhuman vocal learning species such as songbirds and bats (Lipkind et al., 2020; Vernes, Janik, et al., 2021). Infants experiment with utterances without producing proper adult vocalizations. In fact, babbling was long considered a crucial prerequisite for VPL (Fitch, 2006a). However, this evidence comes only from songbirds and humans and thus requires a broader cross-species perspective to address (ter Haar et al., 2021). Vocal development preceding the adult repertoire has been documented in some vocal learning species as well as in species that are considered limited vocal learners (e.g., New World monkeys and giant otters; ter Haar et al., 2021). It is still unknown whether babbling occurs during vocal ontogeny in young elephants or manatees, and if so, to what extent. The increasing diversity of vocal combinations in calves, as well as the snorts that precede correct trumpets, may be promising areas for further exploration based on the current (albeit still limited) understanding of elephant vocal ontogeny. Because of the long developmental periods in elephants and Sirenia, studying their vocal ontogeny is challenging and time-consuming, but observing several mother–calf units across time is a crucial first step toward a better understanding of vocal development and the possible significance of vocal learning in these taxa.

Discussion

To gain insight into the evolution and diversification of vocal learning capacities over time, scholars must investigate various aspects of *Paenungulata* vocal

communication, some of which are described in this article. In the wild, the difficulties associated with being unable to detect or correctly identify calling animals are likely to make it overwhelming to reliably investigate vocal learning in manatees and dugongs. Experimental investigation of vocal learning abilities in trained individuals may be a promising approach to increase the number of species that can be compared. Stoeger and Baotic (2021) and Stoeger et al. (2021) showed that elephants can be trained to vocalize on verbal cue, which revealed the elephants' ability of VUL. For an experimental demonstration of VUL, an animal has to reliably produce a call in response to a specific cue (Shapiro et al., 2004). More convincing evidence appears when an animal remains silent or stops vocalizing in response to different cues. The most complex level of VUL is shown when an animal is capable of emitting different call types in response to distinct cues (Shapiro et al., 2004). Individual African savanna elephants have mastered the most complex level, producing up to seven call types in response to distinct verbal cues (Stoeger & Baotic, 2021). Asian elephants have also been observed to vocalize on verbal cue (Beeck et al., 2021), though further scientific experiments are pending. Because African forest elephants are not kept in captivity and are typically difficult to monitor, we have no data on this species. Similar tests could be conducted with trained Sirenians. In captive settings, experimental explorations of their vocal learning ability are most plausible for the West Indian and Amazonian manatees because of extensive rescue and rehabilitation efforts throughout their range and facilities housing abundant numbers of animals. The understudied African manatee will prove more difficult to study given the lack of facilities for the care of captive or wild rescued animals.

When similar training paradigms are used, it is possible to compare vocal learning capacities directly. Positive reinforcement training, for example, could be applied to reveal plasticity and creativity in sound production. A study using contingency learning found that reinforced variability induced novelty and creativity in sounds and sound production mechanisms in walrus (*Odobenus rosmarus*). However, Vernes, Kriengwatana, et al. (2021) claimed that if reward-driven (food or positive social feedback) modifications result in vocalizations outside of the recognized vocal repertoire, it may theoretically include processes in addition to VPL.

Socially guided VPL is suggested to require additional connections between the VPL and the social motivation system (Syal & Finlay, 2011). However, it is important to note that the effect of social factors, feedback, and reinforcement on VPL in nonhuman animals is relatively

poorly understood (Chen et al., 2016; Theofanopoulou et al., 2017; Williams, 2004). In human infants, social positive feedback is crucial for early speech learning (Goldstein Michael et al., 2003; Kuhl, 2007). The importance of social feedback has also been shown in the song learning of young male zebra finches (Carouso-Peck & Goldstein, 2019) and is well documented for parrots (Pepperberg, 1997); killer whales (Musser et al., 2014); and, to some extent, in Calimero—a male African elephant that was cross-socialized and raised among Asian females and imitated squeaking sounds (Poole et al., 2005). The social environment and social relationships are important in communication, and vocal learning may be influenced by social motivation. In training experiments, the handlers, not conspecifics, are the social partners interacting with the animals. In case of elephants, we know that animals and handlers form social ties that have good consequences on both sides, including operational and affective benefits (Carlstead et al., 2019; Hart, 1994; Rossman et al., 2017). Therefore, in our opinion, positive reinforcement training can be a valuable method in studying aspects of vocal learning in elephants and Sirenians, while taking into consideration the potentially different neuronal mechanisms involved. These include aspects of (a) VUL, such as context variability and timing of call production, and (b) VPL, such as the accuracy with which auditory models are copied (as well as acoustic model divergence); the degree and type of change or alteration; and time in terms of when learning occurs (testing different age and sex classes), how long it takes, and how long learned effects are maintained. To establish a thorough understanding of the vocal learning phenotype and the functional relevance within the natural communication system, approaches addressing the variety in form and function of calls, sound generation variability, and vocal ontogeny of vocalizations of captive and free-ranging animals are required.

Finally, to fully comprehend the evolution of the vocal learning trait in *Paenungulata*, we must assess the original selection advantages that led to its emergence. Currently, broad cross-taxa comparative evidence supports two primary hypotheses (out of several). To begin, the sexual selection hypothesis proposes that VPL evolves to allow expansion of vocal repertoires in response to mating preferences for more complex vocalizations or songs. Second, the information-sharing hypothesis postulates that as vocal repertoires expand in response to kin selection, information sharing between relatives is favored (Nowicki & Searcy, 2014). It is argued that, for VPL to develop, there needs to be a direct benefit for those individuals that modify their vocalizations.

Jarvis (2006) made a similar argument, claiming that VPL was selected for by two factors: mating preference for a variety of vocalizations and the need for rapid adaptation to propagate sound in various environments. Yet Jarvis added an interesting thought: Selection against vocal learning occurred by predation (where varied vocalizations make an easier target). He argued that once predation pressure was overcome, learned vocalizations could have been used for other functions, including complex communication. According to Jarvis, VPLers have few if any predators, or are among the top predators themselves (humans and killer whales). Adult elephants are also relatively unaffected by predators (aside from humans) that occasionally prey on their calves, such as lions, tigers, and crocodiles (Andheria et al., 2007; Loveridge et al., 2006). Also, only a few records exist of large sharks, crocodiles, or killer whales preying on adult manatees and dugongs (Weller, 2009).

In addition, several major constraints might have favored the development of VPL. One particular interesting constraint highlighted by Verpooten (2021) (and initially proposed by Janik & Slater, 1997) is spatial dimensionality of the environment (in combination with sexual selection, because of the difficulty of coordinating and finding mates). At the species level, a great majority of complex VPL's currently inhabit three-dimensional habitats such as air and water (except for humans and elephants, who are terrestrial). Elephants do not vocalize to attract mates but rather fight over the access to females, particularly during musth (Poole et al., 2011). Because of the social life of elephants, the social cohesion function of VPL (Sewall et al., 2016) appears to be more reasonable (at least for the extant species). Yet an alternative explanation for the origin of VPL should not be ruled out, as one could argue that ancient three-dimensionality played a role in elephant evolution (Kellogg et al., 2007). Theoretically, this could be consistent with the mate choice and spatial-dimensional environment origin of complex VPL. According to phylogenetic reconstructions based on net surface charge of myoglobin, the last common ancestor of *Sirenia*, *Elephantidae* and *Hyracoidea*, possessed expert diving abilities previously observed only in cetaceans and pinnipeds (Mirceta et al., 2013). This emphasizes the importance of comparing closely related species to gain a better understanding of potential evolutionary pressures that may have selected for the vocal learning trait in a recognized complex terrestrial vocal learner, the elephant. Ultimately, this might reveal that elephants not only have aquatic living relatives but also have relatives that are (complex) vocal learners.

References

- Amrine-Madsen, H., Koepfli, K.-P., Wayne, R. K., & Springer, M. S. (2003). A new phylogenetic marker, apolipoprotein B, provides compelling evidence for eutherian relationships. *Molecular Phylogenetics and Evolution*, *28*, 225–240. [https://doi.org/10.1016/S1055-7903\(03\)00118-0](https://doi.org/10.1016/S1055-7903(03)00118-0)
- Anderson, P. K. (1997). Shark bay dugongs in summer. I: Lek mating. *Behaviour*, *134*, 433–462. <https://doi.org/10.1163/156853997X00629>
- Anderson, P. K. (2002). Habitat, niche, and evolution of sirenian mating systems. *Journal of Mammalian Evolution*, *9*, 55–98. <https://doi.org/10.1023/A:1021383827946>
- Anderson, P. K., & Barclay, R. M. R. (1995). Acoustic signals of solitary dugongs: Physical characteristics and behavioral correlates. *Journal of Mammalogy*, *76*, 1226–1237. <https://doi.org/10.2307/1382616>
- Andheria, A. P., Karanth, K. U., & Kumar, N. S. (2007). Diet and prey profiles of three sympatric large carnivores in Bandipur Tiger Reserve, India. *Journal of Zoology*, *273*, 169–175. <https://doi.org/10.1111/j.1469-7998.2007.00310.x>
- Ando-Mizobata, N., Ichikawa, K., Arai, N., & Kato, H. (2011). Dugong vocalization in relation to ambient noise. *Proceedings of the 6th International Symposium on SEASTAR2000 and Asian Bio-logging Science (The 10th SEASTAR2000 workshop)*, 55–59.
- Araya-Salas, M., & Wright, T. (2013). Open-ended song learning in a hummingbird. *Biology Letters*, *9*, Article 20130625. <https://doi.org/10.1098/rsbl.2013.0625>
- Archie, E. A., Moss, C. J., & Alberts, S. C. (2006b). The ties that bind: genetic relatedness predicts the fission and fusion of social groups in wild African elephants. *Proceedings of the Royal Society B: Biological Sciences*, *273*, 513–522. <https://doi.org/10.1098/rspb.2005.3361>
- Asher, R. J., Novacek, M. J., & Geisler, J. H. (2003). Relationships of endemic African mammals and their fossil relatives based on morphological and molecular evidence. *Journal of Mammalian Evolution*, *10*, 131–194. <https://doi.org/10.1023/A:1025504124129>
- Beeck, V. C., Heilmann, G., Kerscher, M., & Stoeger, A. S. (2021). A novel theory of Asian elephant high-frequency squeak production. *BMC Biology*, *19*, Article 121. <https://doi.org/10.1186/s12915-021-01026-z>
- Bengtson, J. L. (1981). *Ecology of manatees (Trichechus manatus) in the St Johns River, Florida* [Doctoral dissertation, University of Minnesota].
- Bengtson, J. L., & Fitzgerald, S. M. (1985). Potential role of vocalizations in West Indian manatees. *Journal of Mammalogy*, *66*, 816–819. <https://doi.org/10.2307/1380821>
- Brady, B., Hedwig, D., Trygonis, V., & Gerstein, E. (2020). Classification of Florida manatee (*Trichechus manatus latirostris*) vocalizations. *The Journal of the Acoustical Society of America*, *147*, 1597–1606. <https://doi.org/10.1121/10.0000849>
- Carlstead, K., Paris, S., & Brown, J. L. (2019). Good keeper–elephant relationships in North American zoos are mutually beneficial to welfare. *Applied Animal Behaviour Science*, *211*, 103–111. <https://doi.org/10.1016/j.applanim.2018.11.003>
- Carouso-Peck, S., & Goldstein, M. H. (2019). Female social feedback reveals non-imitative mechanisms of vocal learning in Zebra finches. *Current Biology*, *29*, 631–636. e633. <https://doi.org/10.1016/j.cub.2018.12.026>
- Chapla, M. E., Nowacek, D. P., Rommel, S. A., & Sadler, V. M. (2007). CT scans and 3D reconstructions of Florida manatee (*Trichechus manatus latirostris*) heads and ear bones. *Hearing Research*, *228*, 123–135. <https://doi.org/10.1016/j.heares.2007.01.029>
- Chen, Y., Matheson, L. E., & Sakata, J. T. (2016). Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proceedings of the National Academy of Sciences*, *113*, 6641–6646. <https://doi.org/10.1073/pnas.1522306113>
- Cheney, D. L., & Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences*, *115*, Article 1974. <https://doi.org/10.1073/pnas.1717572115>
- de Silva, S. (2010). Acoustic communication in the Asian elephant, *Elephas maximus maximus*. *Behaviour*, *147*, 825–852. <https://doi.org/10.1163/000579510X495762>
- de Souza, É. M. S., Freitas, L., da Silva Ramos, E. K., Selleghin-Veiga, G., Rachid-Ribeiro, M. C., Silva, F. A., Marmontel, M., dos Santos, F. R., Laudisoit, A., Verheyen, E., Domning, D. P., & Nery, M. F. (2021). The evolutionary history of manatees told by their mitogenomes. *Scientific Reports*, *11*, Article 3564. <https://doi.org/10.1038/s41598-021-82390-2>

- Deutsch, C. J., Reid, J. P., Bonde, R. K., Easton, D. E., Kochman, H. I., & O'Shea, T. J. (2003). Seasonal movements, migratory behavior, and site fidelity of West Indian manatees along the Atlantic Coast of the United States. *Wildlife Monographs*, pp. 1–77. <http://www.jstor.org/stable/3830830>
- Eaton, R. L. (1974). A beluga whale imitates human speech. *Carnivore*, 2, 22–23.
- Eisenberg, J. F., McKay, G. M., & Jainudeen, M. R. (1971). Reproductive behavior of the Asiatic elephant. *Behaviour*, 38, 193–224. <https://doi.org/10.1163/156853971X00087>
- Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, 4, 258–267. [https://doi.org/10.1016/S1364-6613\(00\)01494-7](https://doi.org/10.1016/S1364-6613(00)01494-7)
- Fitch, W. T. (2006a). The biology and evolution of music: A comparative perspective. *Cognition*, 100, 173–215. <https://doi.org/10.1016/j.cognition.2005.11.009>
- Fitch, W. T. (2006b). Production of vocalizations in mammals. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (pp. 115–121). Elsevier. <https://doi.org/10.1016/B0-08-044854-2/00821-X>
- Fitch, W. T. (2010). *The evolution of language*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511817779>
- Fitch, W. T., Neubauer, J., & Herzel, H. (2002). Calls out of chaos: The adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour*, 63, 407–418. <https://doi.org/10.1006/anbe.2001.1912>
- Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 1785–1801. <https://doi.org/10.1098/rstb.2011.0213>
- Goldstein Michael, H., King Andrew, P., & West Meredith, J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences*, 100, 8030–8035. <https://doi.org/10.1073/pnas.1332441100>
- Grossman, C. J., Hamilton, R. E., Wit, M. D., Johnson, J., Faul, R., Herbert, S., Tierney, D., Buot, M., Latham, M. L., Boivin, G., & Boivin, G. P. (2014). The vocalization mechanism of the florida manatee (*Trichechus manatus latirostris*). *Online Journal of Biological Sciences*, 14(2), 127–149. <https://doi.org/10.3844/ojbsci.2014.127.149>
- Hart, L. A. (1994). The Asian elephants–driver partnership: The drivers' perspective. *Applied Animal Behaviour Science*, 40, 297–312. [https://doi.org/10.1016/0168-1591\(94\)90070-1](https://doi.org/10.1016/0168-1591(94)90070-1)
- Hartman, D. S. (1979). *Ecology and behavior of the manatee (Trichechus manatus) in Florida* (No. 5). American Society of Mammalogists. <https://doi.org/10.5962/bhl.title.39474>
- Herbst, C. T., Stoeger, A. S., Frey, R., Lohscheller, J., Titze, I. R., Gumpenberger, M., & Fitch, W. T. (2012). How low can you go? Physical production mechanism of elephant infrasonic vocalizations. *Science*, 337, 595–599. <https://doi.org/10.1126/science.1219712>
- Herler, A., & Stoeger, A. S. (2012). Vocalizations and associated behaviour of Asian elephant (*Elephas maximus*) calves. *Behaviour*, 149, 575–599. <https://doi.org/10.1163/156853912X648516>
- Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, 28, 60–65. <https://doi.org/10.1016/j.conb.2014.06.010>
- Janik, V. M., & Slater, P. J. B. (1997). Vocal learning in mammals. In P. J. B. Slater, J. S. Rosenblatt, C. T. Snowdon, & M. Milinski (Eds.), *Advances in the study of behavior* (Vol. 26, pp. 59–99). Academic Press. [https://doi.org/10.1016/S0065-3454\(08\)60377-0](https://doi.org/10.1016/S0065-3454(08)60377-0)
- Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, 60, 1–11. <https://doi.org/10.1006/anbe.2000.1410>
- Jarvis, E. D. (2006). Selection for and against vocal learning in birds and mammals. *Ornithological Science*, 5, 5–14. <https://doi.org/10.2326/osj.5.5>
- Jarvis, E. D. (2019). Evolution of vocal learning and spoken language. *Science*, 366, 50–54. <https://doi.org/10.1126/science.aax0287>

- Johnson, K. E., & Clark, C. J. (2020). Ontogeny of vocal learning in a hummingbird. *Animal Behaviour*, *167*, 139–150. <https://doi.org/10.1016/j.anbehav.2020.07.010>
- Kellogg, M. E., Burkett, S., Dennis, T. R., Stone, G., Gray, B. A., McGuire, P. M., Zori, R. T., & Stanyon, R. (2007). Chromosome painting in the manatee supports Afrotheria and Paenungulata. *BMC Evolutionary Biology*, *7*, Article 6. <https://doi.org/10.1186/1471-2148-7-6>
- Kershenbaum, A., Ilany, A., Blaustein, L., & Geffen, E. (2012). Syntactic structure and geographical dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 2974–2981. <https://doi.org/10.1098/rspb.2012.0322>
- Kjer, K. M., & Honeycutt, R. L. (2007). Site specific rates of mitochondrial genomes and the phylogeny of eutheria. *BMC Evolutionary Biology*, *7*, Article 8. <https://doi.org/10.1186/1471-2148-7-8>
- Kuhl, P. K. (2007). Is speech learning ‘gated’ by the social brain? *Developmental Science*, *10*, 110–120. <https://doi.org/10.1111/j.1467-7687.2007.00572.x>
- Landrau-Giovanetti, N., Mignucci-Giannoni, A. A., & Reidenberg, J. S. (2014). Acoustical and anatomical determination of sound production and transmission in West Indian (*Trichechus manatus*) and Amazonian (*T. inunguis*) manatees. *The Anatomical Record*, *297*, 1896–1907. <https://doi.org/10.1002/ar.22993>
- Lattenkamp, E. Z. (2020). *Vocal learning in the pale spear-nosed bat, Phyllostomus discolor* [Doctoral dissertation, Radboud University Nijmegen]. <http://hdl.handle.net/21.11116/0000-0005-DF33-7>
- Lattenkamp, E. Z., & Vernes, S. C. (2018). Vocal learning: A language-relevant trait in need of a broad cross-species approach. *Current Opinion in Behavioral Sciences*, *21*, 209–215. <https://doi.org/10.1016/j.cobeha.2018.04.007>
- Lattenkamp, E. Z., Vernes, S. C., & Wiegrebe, L. (2018). Volitional control of social vocalisations and vocal usage learning in bats. *Journal of Experimental Biology*, *221*(14), Article jeb.180729. <https://doi.org/10.1242/jeb.180729>
- Lee, P. C. (1987). Allomothering among African elephants. *Animal Behaviour*, *35*, 278–291. [https://doi.org/10.1016/S0003-3472\(87\)80234-8](https://doi.org/10.1016/S0003-3472(87)80234-8)
- Lee, P. C., & Moss, C. J. (2011). Calf development and maternal rearing strategies. In C. J. Moss, H. Croze, & P. C. Lee (Eds.), *The Amboseli elephants: A long-term perspective on a long-lived mammal* (pp. 225–237). University of Chicago Press. <https://doi.org/10.7208/9780226542263-003>
- Lee, P. C., Poole, J. H., Njiraini, N., & Moss, C. J. (2011). Male social dynamics: independence and beyond. In C. J. Moss, H. J. Croze, & P. C. Lee (Eds.), *The Amboseli elephants: A long-term perspective on a long-lived mammal* (pp. 260–271). University of Chicago Press. <https://doi.org/10.7208/chicago/9780226542263.003.0017>
- Leighty, K. A., Soltis, J., Wesolek, C. M., & Savage, A. (2008). Rumble vocalizations mediate interpartner distance in African elephants, *Loxodonta africana*. *Animal Behaviour*, *76*, 1601–1608. <https://doi.org/10.1016/j.anbehav.2008.06.022>
- 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. <https://doi.org/10.1080/09524622.2003.9753499>
- Lipkind, D., Geambasu, A., & Levelt, C. C. (2020). The development of structured vocalizations in songbirds and humans: A comparative analysis. *Topics in Cognitive Science*, *12*, 894–909. <https://doi.org/10.1111/tops.12414>
- Loveridge, A. J., Hunt, J. E., Murindagomo, F., & Macdonald, D. W. (2006). Influence of drought on predation of elephant (*Loxodonta africana*) calves by lions (*Panthera leo*) in an African wooded savannah. *Journal of Zoology*, *270*, 523–530. <https://doi.org/10.1111/j.1469-7998.2006.00181.x>
- Madsen, P. T., Jensen, F. H., Carder, D., & Ridgway, S. (2012). Dolphin whistles: A functional misnomer revealed by heliox breathing. *Biology Letters*, *8*, 211–213. <https://doi.org/10.1098/rsbl.2011.0701>
- Mann, D. A., O’Shea, T. J., & Nowacek, D. P. (2006). Nonlinear dynamics in manatee vocalizations. *Marine Mammal Science*, *22*, 548–555. <https://doi.org/10.1111/j.1748-7692.2006.00036.x>

- Marsh, H., O'Shea, T. J., & Reynolds, J. E., III. (2011). *Ecology and conservation of the sirenia: Dugongs and manatees*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139013277>
- Martins, P. T., & Boeckx, C. (2020). Vocal learning: Beyond the continuum. *PLoS Biology*, *18*, Article e3000672. <https://doi.org/10.1371/journal.pbio.3000672>
- Matthee, C. A., Eick, G., Willows-Munro, S., Montgelard, C., Pardini, A. T., & Robinson, T. J. (2007). Indel evolution of mammalian introns and the utility of non-coding nuclear markers in eutherian phylogenetics. *Molecular Phylogenetics and Evolution*, *42*, 827–837. <https://doi.org/10.1016/j.ympev.2006.10.002>
- McComb, K., Moss, C., Sayialel, S., & Baker, L. (2000). Unusually extensive networks of vocal recognition in African elephants. *Animal Behaviour*, *59*, 1103–1109. <https://doi.org/10.1006/anbe.2000.1406>
- McComb, K., Reby, D., Baker, L., Moss, C., & Sayialel, S. (2003). Long-distance communication of acoustic cues to social identity in African elephants. *Animal Behaviour*, *65*, 317–329. <https://doi.org/10.1006/anbe.2003.2047>
- Meyer, M., Palkopoulou, E., Baleka, S., Stiller, M., Penkman, K. E. H., Alt, K. W., Ishida, Y., Mania, D., Mallick, S., Meijer, T., Meller, H., Nagel, S., Nickel, B., Ostritz, S., Rohland, N., Schauer, K., Schüler, T., Roca, A. L., Reich, D., Shapiro, B., & Hofreiter, M. (2017). Palaeogenomes of Eurasian straight-tusked elephants challenge the current view of elephant evolution. *eLife*, *6*, Article e25413. <https://doi.org/10.7554/eLife.25413>
- Miksis-Olds, J. L., & Tyack, P. L. (2009). Manatee (*Trichechus manatus*) vocalization usage in relation to environmental noise levels. *The Journal of the Acoustical Society of America*, *125*, 1806–1815. <https://doi.org/10.1121/1.3068455>
- Mirceta, S., Signore, A. V., Burns, J. M., Cossins, A. R., Campbell, K. L., & Berenbrink, M. (2013). Evolution of mammalian diving capacity traced by myoglobin net surface charge. *Science*, *340*, Article 1234192. <https://doi.org/10.1126/science.1234192>
- Moss, C. J., & Lee, P. C. (2011). The Amboseli Elephants: A Long-Term Perspective on a Long-Lived Mammal. In J. M. Cynthia, C. Harvey, & C. L. Phyllis (Eds.), *13. Female Social Dynamics: Fidelity and Flexibility* (pp. 205–223). University of Chicago Press. <https://doi.org/10.7208/9780226542263-015>
- Moss, C. J., & Poole, J. H. (1983). Relationships and social structure of African elephants. In R. A. Hinde (Ed.), *Primate social relations: An integrated approach* (pp. 315–325). Blackwell Scientific.
- Mou Sue, L., Chen, D. H., Bonde, R. K., & O'Shea, T. J. (1990). Distribution and status of manatees (*Trichechus manatus*) in Panama. *Marine Mammal Science*, *6*, 234–241. <https://doi.org/10.1111/j.1748-7692.1990.tb00247.x>
- Musser, W. B., Bowles, A. E., Grebner, D. M., & Crance, J. L. (2014). Differences in acoustic features of vocalizations produced by killer whales cross-socialized with bottlenose dolphins. *The Journal of the Acoustical Society of America*, *136*, 1990–2002. <https://doi.org/10.1121/1.4893906>
- Nair, S., Balakrishnan, R., Seelamantula, C. S., & Sukumar, R. (2009). Vocalizations of wild Asian elephants (*Elephas maximus*): Structural classification and social context. *The Journal of the Acoustical Society of America*, *126*, 2768–2778. <https://doi.org/10.1121/1.3224717>
- Nielsen, M. L. K., Bejder, L., Videsen, S. K. A., Christiansen, F., & Madsen, P. T. (2019). Acoustic crypsis in southern right whale mother–calf pairs: Infrequent, low-output calls to avoid predation? *Journal of Experimental Biology*, *222*(13), Article jeb.190728. <https://doi.org/10.1242/jeb.190728>
- Nishihara, H., Satta, Y., Nikaido, M., Thewissen, J. G. M., Stanhope, M. J., & Okada, N. (2005). A retroposon analysis of Afrotherian phylogeny. *Molecular Biology and Evolution*, *22*, 1823–1833. <https://doi.org/10.1093/molbev/msi179>
- Noll, J. D. (1983). The origin of the term “harelip.” *Cleft Palate Journal*, *20*, 169–171.
- Nottebohm, F. (1972). The origins of vocal learning. *The American Naturalist*, *106*, 116–140. <https://doi.org/10.1086/282756>

- Nowacek, D. P., Casper, B. M., Wells, R. S., Nowacek, S. M., & Mann, D. A. (2003). Intraspecific and geographic variation of West Indian manatee (*Trichechus manatus spp.*) vocalizations (L). *The Journal of the Acoustical Society of America*, *114*, 66–69. <https://doi.org/10.1121/1.1582862>
- Nowicki, S., & Searcy, W. A. (2014). The evolution of vocal learning. *Current Opinion in Neurobiology*, *28*, 48–53. <https://doi.org/10.1016/j.conb.2014.06.007>
- O'Shea, T. J., & Poché, L. B., Jr. (2006). Aspects of underwater sound communication in Florida manatees (*Trichechus manatus latirostris*). *Journal of Mammalogy*, *87*, 1061–1071. <https://doi.org/10.1644/06-MAMM-A-066R1.1>
- Pardini, A. T., O'Brien, P. C. M., Fu, B., Bonde, R. K., Elder, F. F. B., Ferguson-Smith, M. A., Yang, F., & Robinson, T. J. (2007). Chromosome painting among proboscidea, hyracoidea and sirenia: Support for Paenungulata (*Afrotheria*, *Mammalia*) but not Tethytheria. *Proceedings of the Royal Society B: Biological Sciences*, *274*, 1333–1340. <https://doi.org/10.1098/rspb.2007.0088>
- Pardo, M. A., Poole, J. H., Stoeger, A. S., Wrege, P. H., O'Connell-Rodwell, C. E., Padmalal, U. K., & de Silva, S. (2019). Differences in combinatorial calls among the 3 elephant species cannot be explained by phylogeny. *Behavioral Ecology*, *30*, 809–820. <https://doi.org/10.1093/beheco/arz018>
- Parks, S. E., Cusano, D. A., Van Parijs, S. M., & Nowacek, D. P. (2019). Acoustic crypsis in communication by North Atlantic right whale mother–calf pairs on the calving grounds. *Biology Letters*, *15*, Article 20190485. <https://doi.org/10.1098/rsbl.2019.0485>
- Payne, K. B., Langbauer, W. R., & Thomas, E. M. (1986). Infrasonic calls of the Asian elephant (*Elephas maximus*). *Behavioral Ecology and Sociobiology*, *18*, 297–301. <https://doi.org/10.1007/BF00300007>
- Pepperberg, I. M. (1997). Social influences on the acquisition of human-based codes in parrots and nonhuman primates. In C. T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development* (pp. 157–177). Cambridge University Press. <https://doi.org/10.1017/CBO9780511758843.009>
- Pepperberg, I. M. (2010). Vocal learning in Grey parrots: A brief review of perception, production, and cross-species comparisons. *Brain and Language*, *115*, 81–91. <https://doi.org/10.1016/j.bandl.2009.11.002>
- Petkov, C., & Jarvis, E. (2012). Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, *4*, Article 12. <https://doi.org/10.3389/fnevo.2012.00012>
- Phillips, R., Niezrecki, C., & Beusse, D. O. (2003). Determination of West Indian manatee vocalization levels and rate. *The Journal of the Acoustical Society of America*, *115*, 422–428. <https://doi.org/10.1121/1.1635839>
- Poole, J. H. (1989). Mate guarding, reproductive success and female choice in African elephants. *Animal Behaviour*, *37*, 842–849. [https://doi.org/10.1016/0003-3472\(89\)90068-7](https://doi.org/10.1016/0003-3472(89)90068-7)
- Poole, J. H. (2011). Behavioral contexts of elephant acoustic communication. In C. J. Moss, H. Croze, & P. C. Lee (Eds.), *The Amboseli elephants: A long-term perspective on a long-lived mammal* (pp. 125–161). The University of Chicago. <https://doi.org/10.7208/chicago/9780226542263.003.0009>
- Poole, J. H., Lee, P. C., Njiraini, N., & Moss, C. J. (2011). Longevity, competition, and musth: A long-term perspective on male reproductive strategies. In C. J. Moss, H. Croze, & P. C. Lee (Eds.), *The Amboseli elephants: A long-term perspective on a long-lived mammal* (pp. 272–288). University of Chicago Press. <https://doi.org/10.7208/9780226542263-020>
- Poole, J. H., Payne, K., Langbauer, W. R., & Moss, C. J. (1988). The social contexts of some very low frequency calls of African elephants. *Behavioral Ecology and Sociobiology*, *22*, 385–392. <https://doi.org/10.1007/bf00294975>
- Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., & Watwood, S. (2005). Elephants are capable of vocal learning. *Nature*, *434*, 455–456. <https://doi.org/10.1038/434455a>
- Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology*, *63*, 1050–1056. <https://doi.org/10.1139/z85-157>

- Ramos, E. A., Maust-Mohl, M., Collom, K. A., Brady, B., Gerstein, E. R., Magnasco, M. O., & Reiss, D. (2020). The Antillean manatee produces broadband vocalizations with ultrasonic frequencies. *The Journal of the Acoustical Society of America*, *147*, EL80–EL86. <https://doi.org/10.1121/10.0000602>
- Ravnani, A., & Garcia, M. (2021). A cross-species framework to identify vocal learning abilities in mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*, Article 20200394. <https://doi.org/10.1098/rstb.2020.0394>
- Reichmuth, C., & Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. *Current Opinion in Neurobiology*, *28*, 66–71. <https://doi.org/10.1016/j.conb.2014.06.011>
- Reynolds, J. E., III, & Wilcox, J. R. (1994). Observations of Florida manatees (*Trichechus manatus latirostris*) around selected power plants in winter. *Marine Mammal Science*, *10*, 163–177. <https://doi.org/10.1111/j.1748-7692.1994.tb00258.x>
- Reynolds, J. E., III. (1981). Aspects of the social behaviour and herd structure of a semi-isolated colony of West Indian manatees, *Trichechus manatus*. *Mammalia*, *45*(4), 431–452. <https://doi.org/10.1515/mamm.1981.45.4.431>
- Rommel, S. A., Pabst, D. A., McLellan, W. A., & Costidis, A. M. (2018). Skull. In B. Würsig, J. G. M. Thewissen, & K. M. Kovacs (Eds.), *Encyclopedia of marine mammals* (3rd ed., pp. 871–882). Academic Press. <https://doi.org/10.1016/B978-0-12-804327-1.00232-6>
- Rossmann, Z. T., Padfield, C., Young, D., & Hart, L. A. (2017). Elephant-initiated interactions with humans: Individual differences and specific preferences in captive African elephants (*Loxodonta africana*). *Frontiers in Veterinary Science*, *4*. <https://doi.org/10.3389/fvets.2017.00060>
- Rycyk, A. M., Factheu, C., Ramos, E. A., Brady, B. B., Kikutchi, M., Nations, H. F., Kapfer, K., Hampton, C. M., Garcia, E. R., & Kamla, A. (2021). First characterization of vocalizations and passive acoustic monitoring of the vulnerable African manatee (*Trichechus senegalensis*). *The Journal of the Acoustical Society of America*, (in press). <https://doi.org/10.1121/10.0006734>
- Schusterman, R. J., & Reichmuth, C. (2008). Novel sound production through contingency learning in the Pacific walrus (*Odobenus rosmarus divergens*). *Animal Cognition*, *11*, 319–327. <https://doi.org/10.1007/s10071-007-0120-5>
- Seiffert, E. R. (2007). A new estimate of afrotherian phylogeny based on simultaneous analysis of genomic, morphological, and fossil evidence. *BMC Evolutionary Biology*, *7*, Article 224. <https://doi.org/10.1186/1471-2148-7-224>
- Sewall, K. B., Young, A. M., & Wright, T. F. (2016). Social calls provide novel insights into the evolution of vocal learning. *Animal Behaviour*, *120*(1), 163–172. <https://doi.org/10.1016/j.anbehav.2016.07.031>
- Shapiro, A. D., Slater, P. J. B., & Janik, V. M. (2004). Call usage learning in gray seals (*Halichoerus grypus*). *Journal of Comparative Psychology*, *118*, 447–454. <https://doi.org/10.1037/0735-7036.118.4.447>
- Soltis, J. (2010). Vocal communication in African elephants (*Loxodonta africana*). *Zoo Biology*, *29*, 192–209. <https://doi.org/10.1002/zoo.20251>
- Sousa-Lima, R. S., Paglia, A. P., & Da Fonseca, G. A. B. (2002). Signature information and individual recognition in the isolation calls of Amazonian manatees, *Trichechus inunguis* (Mammalia: Sirenia). *Animal Behaviour*, *63*, 301–310. <https://doi.org/10.1006/anbe.2001.1873>
- Sousa-Lima, R. S., Paglia, A. P., & Da Fonseca, G. A. B. (2008). Gender, age, and identity in the isolation calls of Antillean Manatees (*Trichechus manatus manatus*). *Aquatic Mammals*, *34*, 109–122. <https://doi.org/10.1578/AM.34.1.2008.109>
- Springer, M. S., & Murphy, W. J. (2007). Mammalian evolution and biomedicine: New views from phylogeny. *Biological Reviews*, *82*, 375–392. <https://doi.org/10.1111/j.1469-185X.2007.00016.x>

- Springer, M. S., Signore, A. V., Pajmans, J. L. A., Vélez-Juarbe, J., Domning, D. P., Bauer, C. E., He, K., Crerar, L., Campos, P. F., Murphy, W. J., Meredith, R. W., Gatesy, J., Willerslev, E., MacPhee, R. D. E., Hofreiter, M., & Campbell, K. L. (2015). Interordinal gene capture, the phylogenetic position of Steller's sea cow based on molecular and morphological data, and the macroevolutionary history of Sirenia. *Molecular Phylogenetics and Evolution*, *91*, 178–193. <https://doi.org/10.1016/j.ympev.2015.05.022>
- Steel, C. (1982). *Vocalization patterns and corresponding behavior of the West Indian manatee (Trichechus manatus)*. Florida Institute of Technology, Melbourne.
- Stoeger, A. S., & Baotic, A. (2021). Operant control and call usage learning in African elephants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*, Article 20200254. <https://doi.org/10.1098/rstb.2020.0254>
- Stoeger, A. S., Baotic, A., & Heilmann, G. (2021). Vocal creativity in elephant sound production. *Biology*, *10*. <https://doi.org/10.3390/biology10080750>
- 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*, 1700–1710. <https://doi.org/10.1121/1.3605538>
- Stoeger, A. S., & de Silva, S. (2014). African and Asian elephant vocal communication: A cross-species comparison. In G. Witzany (Ed.), *Biocommunication of animals* (pp. 21–39). Springer. https://doi.org/10.1007/978-94-007-7414-8_3
- Stoeger, A. S., Heilmann, G., Zeppelzauer, M., Ganswindt, A., Hensman, S., & Charlton, B. D. (2012). Visualizing sound emission of elephant vocalizations: Evidence for two rumble production types. *PLoS ONE*, *7*, Article e48907. <https://doi.org/10.1371/journal.pone.0048907>
- Stoeger, A. S., & Manger, P. (2014). Vocal learning in elephants: Neural bases and adaptive context. *Current Opinion in Neurobiology*, *28*, 101–107. <https://doi.org/10.1016/j.conb.2014.07.001>
- Stoeger, A. S., Mietchen, D., Oh, S., de Silva, S., Herbst, Christian T., Kwon, S., & Fitch, W. T. (2012). An Asian Elephant Imitates Human Speech. *Current Biology*, *22*, 2144–2148. <https://doi.org/10.1016/j.cub.2012.09.022>
- 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*, 3922–3931. <https://doi.org/10.1121/1.2722216>
- Syal, S., & Finlay, B. L. (2011). Thinking outside the cortex: social motivation in the evolution and development of language. *Developmental Science*, *14*, 417–430. <https://doi.org/10.1111/j.1467-7687.2010.00997.x>
- Tassy, P., & Shoshani, J. (1988). The Tethytheria: Elephants and their relatives. In M. J. Benton (Ed.), *The phylogeny and classification of the tetrapods* (Vol. 2 Mammals, pp. 283–315). Clarendon Press.
- ter Haar, S. M., Fernandez, A. A., Gratier, M., Knörnschild, M., Levelt, C., Moore, R. K., Vellema, M., Wang, X., & Oller, D. K. (2021). Cross-species parallels in babbling: Animals and algorithms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*, Article 20200239. <https://doi.org/10.1098/rstb.2020.0239>
- Theofanopoulou, C., Boeckx, C., & Jarvis, E. D. (2017). A hypothesis on a role of oxytocin in the social mechanisms of speech and vocal learning. *Proceedings of the Royal Society B: Biological Sciences*, *284*, Article 20170988. <https://doi.org/10.1098/rspb.2017.0988>
- Turkalo, A., & Fay, J. M. (1995). Studying forest elephants by direct observation: preliminary results from the Dzanga clearing, Central African Republic. *Pachyderm*, *20*, 45–54.
- Turkalo, A. K., Wrege, P. H., & Wittemyer, G. (2013). Long-Term Monitoring of Dzanga Bai Forest Elephants: Forest Clearing Use Patterns. *PLoS ONE*, *8*, e85154. <https://doi.org/10.1371/journal.pone.0085154>
- Tyack, P. L. (2020). A taxonomy for vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*, Article 20180406. <https://doi.org/10.1098/rstb.2018.0406>
- Umeed, R., Niemeyer Attademo, F. L., & Bezerra, B. (2018). The influence of age and sex on the vocal repertoire of the Antillean manatee (*Trichechus manatus manatus*) and their responses to call playback. *Marine Mammal Science*, *34*, 577–594. <https://doi.org/10.1111/mms.12467>

- van Dijk, M. A. M., Madsen, O., Catzeflis, F., Stanhope, M. J., de Jong, W. W., & Pagel, M. (2001). Protein sequence signatures support the African clade of mammals. *Proceedings of the National Academy of Sciences*, *98*, Article 188. <https://doi.org/10.1073/pnas.98.1.188>
- Vernes, S. C., Janik, V. M., Fitch, W. T., & Slater, P. J. B. (2021). Vocal learning in animals and humans. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*, Article 20200234. <https://doi.org/10.1098/rstb.2020.0234>
- Vernes, S. C., Kriengwatana, B. P., Beeck, V. C., Fischer, J., Tyack, P. L., ten Cate, C., & Janik, V. M. (2021). The multi-dimensional nature of vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*, Article 20200236. <https://doi.org/10.1098/rstb.2020.0236>
- Verpooten, J. (2021). Complex vocal learning and three-dimensional mating environments. *Biology & Philosophy*, *36*, Article 12. <https://doi.org/10.1007/s10539-021-09786-2>
- Videsen, S. K. A., Bejder, L., Johnson, M., & Madsen, P. T. (2017). High suckling rates and acoustic crypsis of humpback whale neonates maximise potential for mother–calf energy transfer. *Functional Ecology*, *31*, 1561–1573. <https://doi.org/10.1111/1365-2435.12871>
- Vidya, T. N. C. (2014). Novel behaviour shown by an Asian elephant in the context of allomothering. *Acta Ethologica*, *17*, 123–127. <https://doi.org/10.1007/s10211-013-0168-y>
- Vidya, T. N. C., & Sukumar, R. (2005). Social organization of the Asian elephant (*Elephas maximus*) in southern India inferred from microsatellite DNA. *Journal of Ethology*, *23*, 205–210. <https://doi.org/10.1007/s10164-005-0144-8>
- Weller, D. W. (2009). Predation on marine mammals. In W. F. Perrin, B. Würsig, & J. G. M. Thewissen (Eds.), *Encyclopedia of marine mammals* (2nd ed., pp. 923–932). Academic Press. <https://doi.org/10.1016/B978-0-12-373553-9.00210-8>
- Wells, R., Boness, D., & Rathbun, G. (1999). Behavior. In *Biology of marine mammals* (pp. 324–422). Smithsonian Institution Press.
- Williams, H. (2004). Birdsong and singing behavior. *Annals of the New York Academy of Sciences*, *1016*, 1–30. <https://doi.org/10.1196/annals.1298.029>
- Wirsing, A. J., Heithaus, M. R., & Dill, L. M. (2007a). Fear factor: Do dugongs (*Dugong dugon*) trade food for safety from tiger sharks (*Galeocerdo cuvier*)? *Oecologia*, *153*, 1031–1040. <https://doi.org/10.1007/s00442-007-0802-3>
- Wirsing, A. J., Heithaus, M. R., & Dill, L. M. (2007b). Living on the edge: Dugongs prefer to forage in microhabitats that allow escape from rather than avoidance of predators. *Animal Behaviour*, *74*, 93–101. <https://doi.org/10.1016/j.anbehav.2006.11.016>
- Wirthlin, M., Chang, E. F., Knörnschild, M., Krubitzer, L. A., Mello, C. V., Miller, C. T., Pfenning, A. R., Vernes, S. C., Tchernichovski, O., & Yartsev, M. M. (2019). A modular approach to vocal learning: Disentangling the diversity of a complex behavioral trait. *Neuron*, *104*, 87–99. <https://doi.org/10.1016/j.neuron.2019.09.036>
- Wittemyer, G., Douglas-Hamilton, I., & Getz, W. M. (2005). The socioecology of elephants: analysis of the processes creating multitiered social structures. *Animal Behaviour*, *69*, 1357–1371. <https://doi.org/10.1016/j.anbehav.2004.08.018>