Acoustics Today

ASA
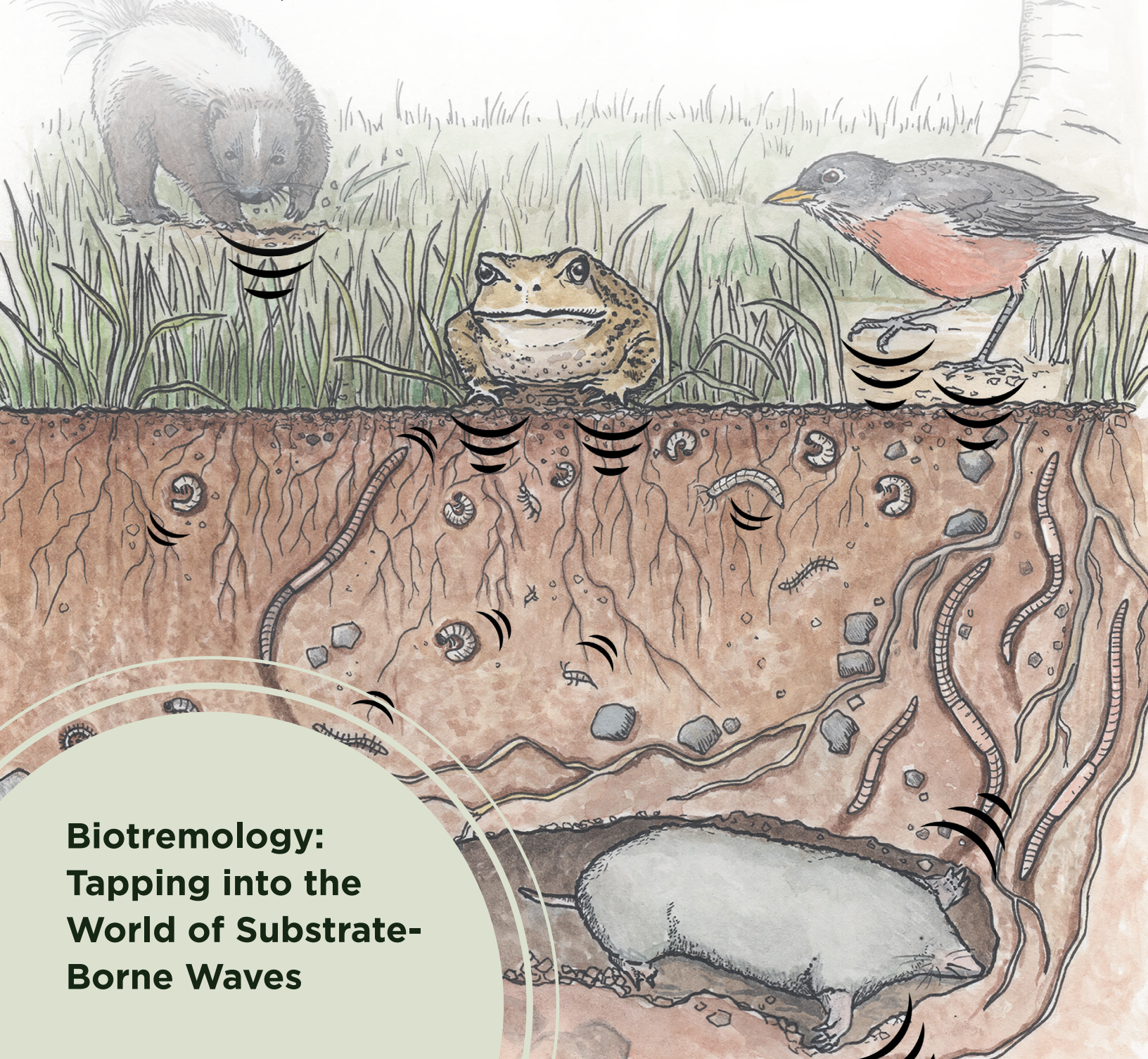ACOUSTICAL SOCIETY
OF AMERICA

An Acoustical Society of America publication

**Biotremology:
Tapping into the
World of Substrate-
Borne Waves**

Acoustics Today

About the Cover

Cross section of a North American lawn, with vibrational users and producers. Cover image from "Biotremology: Tapping into the World of Substrate-Borne Waves" by Louise Roberts and Kyle Wickings (page 49). Paintings by Ann Sanderson, © 2022, Ann Sanderson, all rights reserved; see www.annsciart.com.

From the Editor

Arthur N. Popper

We try to have wide-ranging content in *Acoustics Today* (*AT*), with the goal of having material that all members of the Acoustical Society of America (ASA) will find interesting, and from which they might learn about interesting things outside of their specialized areas. This issue is no different both in terms of species (fishes, humans, insects) and topics covered, although the brain is a part of several of the articles.

Past issues of *AT* have covered sound production mechanisms in many species (see bit.ly/ATC-Bioacoustics) but never in fishes. In our first article, Michael L. Fine and Eric Parmentier write about this fish sound production. Interestingly, acoustic communication arose early in the evolution of fishes (and thus of vertebrates). While many readers are familiar with sound production mechanisms that involve the movement of air (e.g., in mammals), the mechanisms involved in fish sound production are fascinating and quite different, and often involve a variety of different body parts.

In the second article, Steven Greenberg talks about speech but in terms of its rhythm and processing in the brain. At the same time, Steve points out that rhythm is not unique to humans but, in fact, is found throughout the animal kingdom, although it is especially important in human speech.

This is followed in an article by Bonnie K. Lau who talks about pitch perception, a critical part of speech, and its development in infants. Bonnie is particularly interested in how pitch is perceived by young children even before many of the auditory regions of the brain are fully developed.

In the fourth article, Jonathan E. Peelle and Arthur Wingfield provide insight into how the human brain processes speech. An important consideration of this article is how the brain deals with masking noise that could impair the detection of speech. And, as part of the article, Jonathan and Arthur give a lovely "primer" on the human brain and where it processes speech.

The fifth article, by Louise Roberts and Kyle Wickings, introduces a new word to most readers, biotremology. Biotremology is the study of how signals are detected and used by both terrestrial and aquatic animals. In fact, many members of the ASA are familiar with the physical basis of biotremology, substrate vibration. In their article, Louise and Kyle give a broad overview of the topic and show how these signals have great importance to animals ranging in size from elephants to insects.

Our last article brings us to music. After reading this article on guitar sounds by Jesús Alejandro Torres, I can no longer listen to a guitar without thinking not only about the music I'm hearing but also how the music is produced. Jesús introduces us to both acoustic and electric guitars and the fascinating differences between them not only in structure but also in how they produce sounds and the sounds they produce. This article is one of a number we have had over the years about musical instruments, the complete collection of which can be found in "*AT* Collections" at bit.ly/ATC-Music.

As usual, there are several interesting essays in our "Sound Perspectives" section. Our new series, "Conversation with a Colleague" (CwC; see bit.ly/ATC-CWC), features Andone Lavery, an acoustic oceanographer well-known to many members of the ASA. In her article, Andone focuses on her fascinating work in acoustics in oceanography and about how she evolved her current interests and pursuits.

As an aside, Micheal Dent, my colleague who edits the CwC essays, will continue to reach out to technical committee chairs to get suggestions for possible people to include in the series. Our focus is on midcareer scholars, and we are hoping to include essays by people in a variety of different career paths. Indeed, the next CwC will be by a professional architect.

XL3

ACOUSTIC ANALYZER

The perfect solution for acoustic measurements and noise monitoring

.NTi AUDIO 🇨🇭

We annually have an essay from the active and important ASA Student Council (see bit.ly/ATC-Students). In this issue, Megan S. Anderson, Zane Rusk, Colby Cushing, Lucy Ruoqian Cheng, Hilary Kates Varghese, Mark Langhirt, and Elizabeth Weidner tell stories about their research as graduate students. It is great to see that our students are starting out with strong and interesting research programs. If they are, as I suspect, representative of the newer acousticians in the ASA, the Society has a great future not only for its membership but also for its contributions to the overall field.

Other potential acousticians are featured in the piece by Laurie M. Heller about the ASA participation in the International Science and Engineering Fair. However, these prize-winning students are in high school, and they, as demonstrated in the essay, are already doing amazing research in various aspects of acoustics. I propose the suggestion that any acoustics-related undergraduate programs try and "recruit" these remarkable young people.

The ASA is, as many readers know, looking for ways to educate members and the community about acoustics issues. Indeed, ASA members are increasingly engaging in improving communication in both the Society and the community, and *AT* has featured several articles and essays on this topic (see bit.ly/ATC-Communications). In this issue, Kathleen J. Vigness-Raposa, Holly Morin, Christopher Knowlton, and Gail Scowcroft discuss best practices for developing and using education tools online. They base this essay on 20 years of experience with online learning, including webinars, in the Discovery of Sound in the Sea (DOSITS) project. The lessons learned by DOSITS are highly applicable to the ASA and to all members of our community, and I encourage you to review what the authors share.

As you can see, I included several links to "*AT* Collections" in this column (see bit.ly/AT-Collections). "*AT* Collections" is a relatively new feature of the *AT* website that is intended to bring together the more than 325 articles published to date into groupings that apply to a variety of topics, many of which can be used in college courses or for other educational purposes. We welcome new collections based on courses members teach, research interests, job function, or
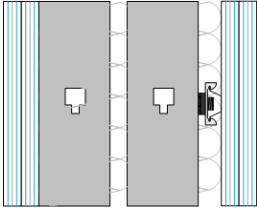
any other purpose. Developing a collection is quite simple. If any reader would like to develop a collection, please drop me an email (apopper@umd.edu) with your ideas.

From the President

Peggy Nelson

It's a great pleasure to write this column as the new Acoustical Society of America (ASA) President. The ASA has meant a great deal to me both professionally and personally. I have been a member of the Society since I was a graduate student, proudly publishing my dissertation work on psychoacoustics and speech recognition by listeners with hearing loss in *The Journal of the Acoustical Society of America* (*JASA*). Since then, I have served on the Standards and Women in Acoustics Committees and recently as vice president. I'm committed to supporting the Society so that it can support others as it has supported me.

As I write my first From the President column, we have just concluded the Denver ASA meeting, only our second in-person meeting since 2020. One of my first tasks was to serve as one of the hosts of the acoustics school for Summer Undergraduate Research or Industry Experience in Acoustics (SURIEA) students (see acousticalsociety.org/suriea), which also took place in person. My experiences at both the meeting and the SURIEA school brought to focus the optimism that I sense about our Society. I hope that you will take a moment to read this, reflect on a few issues facing the ASA, and perhaps reach out to let me know your thoughts. Here are the fundamental strengths that I see in our Society.

- First and foremost, we remain an *exemplary* scientific society, promoting the best of the science and practice of acoustics. In Denver, I attended exceptional talks and viewed outstanding posters. Every month, I read our journals and know that I will find the best acoustics work there. Now between meetings, I continue to enjoy the webinars and journal clubs that also promote the best of our discipline. I'm proud to present and publish my own best work here in the Society, and I hope you all feel the same. We excel at promoting the best in both acoustic science and practice.
- Next, we are *resilient* and well prepared in the face of some formidable recent challenges. It is a clear understatement to describe the past few years as extremely challenging. We have weathered those challenges and

(crossing my fingers) are coming out on the "other side" of the pandemic. How are we able to do so? We are building on the strengths of the generations before us. Leaders of the past have brought us well into the twenty-first century and have left us in a position of strength. Recent leaders (in particular, Past Presidents Vic Sparrow, Diane Kewley-Port, and Maureen Stone) have steered us through the months dominated by the global pandemic. The staff at headquarters has always been exceptional and has faced the challenges of the past few years with creativity and hard work. We've experienced some financial shortfalls, as have many societies, but we're emerging now in a position of good health. Careful stewardship over the next few years will be critical, but we have terrific optimism that we will be solid for generations to come.

- Third, we are *embracing change* and helping to lead the way toward more inclusivity in science. The past two years have seen upheaval and great discord in the world. Within the ASA, we have faced potentially divisive issues, with strong statements backed by action. The ASA is seeing the emergence of early-career members who are leading the way, and this is the key to long-term growth of the Society into its next century. (See the column by immediate Past President Maureen Stone about the approaching 100th anniversary of the ASA at bit.ly/ASA-100.) Many members met the students from our inaugural SURIEA cohort at the Denver meeting. These students are truly impressive and give us great hope for the Society's future. After considering applications from more than a hundred undergraduate students, our second SURIEA cohort members were similarly, but uniquely, impressive. Their talent and excitement for acoustics is further evidence that ASA is benefiting from meeting the challenge of racial equity and justice with bold action. The program will continue, and we hope that next summer even more ASA members will apply to become SURIEA mentors. Through the ASA Committee to Improve Racial Diversity and Inclusion (CIRDI) (see tinyurl.com/3p8hecdm) and the outstanding leadership of Tyrone Porter, we are embracing inclusion and challenging bias. The ASA Executive Council recently

voted to established long-term funding to continue our efforts in these areas. Although there is a long way to go, I am thrilled that we are on a clear path to improving diversity and equity.

These trends leave me hopeful and excited to help lead the Society over the coming year. New challenges will most certainly arise, and some will no doubt surprise us. We have some big decisions to make about future meetings and the role of virtual technology. We are excited to try special sessions in hybrid format at the next two meetings, and we will look forward to getting your feedback on that endeavor.

We must also tighten our fiscal belts in the near term so that we leave the Society just as strong for the next decades. We're still experiencing annual deficits and we are challenged to rein those in. Immediate Past President Maureen Stone has guided us to more fiscal responsibility and to turn our large ship toward a balanced annual budget. Nevertheless, there

is also good fiscal news. The Society has in reserve enough funds to maintain operations for several years, much more than is seen in many organizations. The Acoustical Society Support Foundation (see bit.ly/3we52wg) under the leadership of Jim Miller has continued to grow and expand its programs for grants, fellowships, and opportunities to support the Society's goals. In important news, we will now be able to accept donations directly to these marvelous programs from you, our members.

I hope you share my optimism and that you will continue to invest in our Society by maintaining your membership, coming to meetings if you are able, and donating to the causes that mean the most to you. Already, we are preparing for the next century of leading the discipline of acoustics. I'm grateful to be a part of that and I welcome you to join me. I hope to see you on a webinar, at a committee meeting, and again in Nashville, Tennessee, in December. Let me know (peggynelson@umn.edu) what you think we can do together.

Fish Sound Production: The Swim Bladder

Michael L. Fine and Eric Parmentier

Sound Production in Fishes

Sound production is nearly ubiquitous among various vertebrate taxa from fishes to mammals. However, although terrestrial vertebrates and marine mammals produce sounds that require air movement, fishes are unique in that they use a variety of mechanical systems that typically do not involve air movement. Indeed, the first vertebrates to produce sounds were likely primitive fishes, and the mechanisms of sound production had to be "reinvented" when vertebrates colonized land because fish sound-producing mechanisms likely did not produce enough acoustic energy to carry substantial distances in the low-density air environment (Gharhamani et al., 2014).

In fact, there is not "one" but many sound-producing mechanisms among the over 30,000 extant fish species (Rice et al., 2022). These range from high-speed contraction of muscles that vibrate a gas-filled chamber to stridulation or the rubbing of bones against one another (Fine and Parmentier, 2015).

Perhaps the best understood mechanisms for sound production in fishes involve the subject of this article, the swim bladder, a gas-filled chamber in the abdominal cavity of most bony fishes (sharks and rays do not have a swim bladder!) (**Figure 1**). The primary role of the swim bladder is to control buoyancy so that fishes do not expend energy to maintain their vertical position in the water column (Pelster, 2021). However, in many species, the swim bladder has evolved for other functions including respiration (Pelster, 2021), hearing (Popper et al., 2021), and sound production (Parmentier and Fine, 2016).

Fish Sounds

Since the days of Aristotle, it has been recognized that fishes make sounds. Indeed, studies over decades
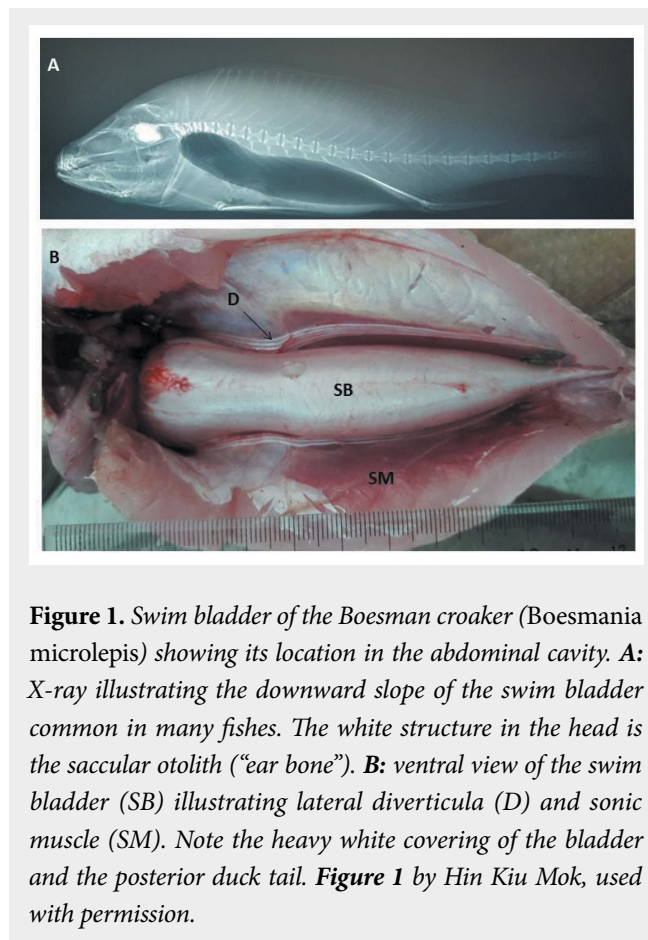
Figure 1. *Swim bladder of the Boesman croaker (*Boesmania microlepis*) showing its location in the abdominal cavity. **A:** X-ray illustrating the downward slope of the swim bladder common in many fishes. The white structure in the head is the saccular otolith ("ear bone"). **B:** ventral view of the swim bladder (SB) illustrating lateral diverticula (D) and sonic muscle (SM). Note the heavy white covering of the bladder and the posterior duck tail. **Figure 1** by Hin Kiu Mok, used with permission.*

demonstrate that a substantial number of fish species produce sounds during courtship, disturbance, aggressive, and other contexts. Examples of fish sounds can be found at DOSITS.org (Discovery of the Sound in the Sea; see tinyurl.com/2s3ahwdp; fishsounds.net). Most fish sounds are a series of pulses that vary in frequency spectra and duration and are emitted with different temporal patterns (e.g., Winn, 1964). Some longer duration sounds can also have a well-developed harmonic structure. Moreover, many fish sounds are species specific, allowing potential call identification as with bird songs (**Figure 2**).
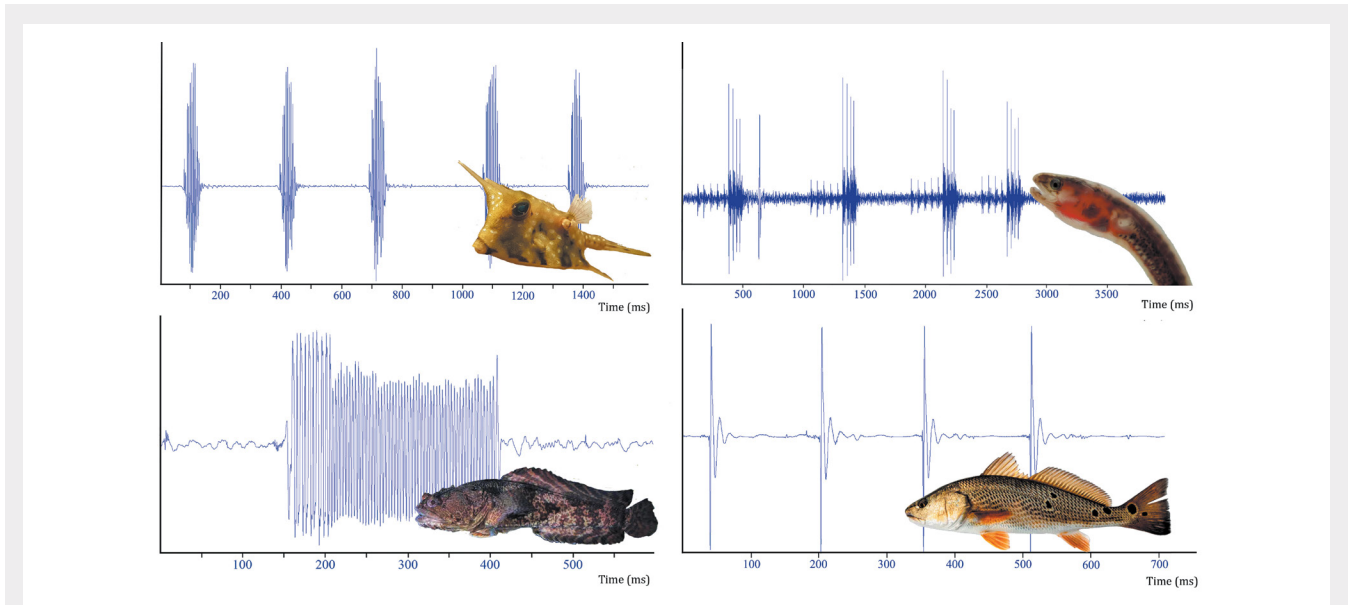
Figure 2. *Oscillograms of sounds from four fish species.* **Top left:** *longhorn cowfish* Lactoria cornuta. **Top right:** *red pearlfish* Encheliophis chardewalli. **Bottom right:** *red drum* Scieanops ocellatus *(see* tinyurl.com/4yhcrec4*)*. **Bottom left:** *oyster toadfish* Opsanus tau *(see* tinyurl.com/4j4rz4nk*)*.

Most fish sounds are generated by stridulation of bony structures or with the swim bladder (Parmentier and Fine, 2016). Stridulation sounds are typically higher pitched than swim bladder sounds and can be produced, for example, by rubbing a process on the base of the pectoral spine in catfish against the pectoral girdle (**Figure 3**) (Fine et al., 1997) or rubbing upper against lower pharyngeal teeth (toothlike structures located in the back of the mouth in fishes) (Bertucci et al., 2014).

Swim bladder sounds are typically generated by extremely fast sonic muscles that excite bladder vibrations (Fine and Parmentier, 2015). For example, a courtship boat whistle call of the oyster toadfish (*Opsanus tau*) with a fundamental frequency of 200 Hz would be stimulated by sonic muscle contractions at 200 Hz (see dosits.org/Oyster-Toadfish). Compared with a bird syrinx or mammalian larynx, the swim bladder, typically a closed structure, has little ability to vary sound frequency. A drop-off of several hertz during a tonal sound is due to muscle fatigue (Mitchell et al., 2008) rather than the ability to regulate tension as on a laryngeal membrane.

Sound production has evolved independently in various fish groups (Fine and Parmentier, 2015). Nearly two-thirds of fish species from over 100 families are known to produce

sounds, although future studies will likely identify many more sound-producing species (Looby et al., 2022). There is still much to be learned about fish sound production and acoustic behavior! Interestingly, the ability to produce sounds has evolved independently at least over 30 times (Parmentier et al., 2021a; Rice et al., 2022). Fish sonic structures

Figure 3. A: *pectoral spine of the channel catfish* Ictalurus punctatus, *illustrating stridulatory surfaces. Spine base is to the right and the dorsal process points upward.* **B:** *ridged undersurface of the process rubs against a rough surface on the cleithrum (a bone in the pectoral girdle)* **(C)**. **Inset:** *magnified rough-rubbing surface that catches the ridges using a stick-slip mechanism similar to a moving bow on a violin (Mohajer et al., 2015). Scale bar, 1 mm in* **A**; *0.5 mm in* **B** *and* **C**. *Modified from Fine et al. (1998), used with permission.*
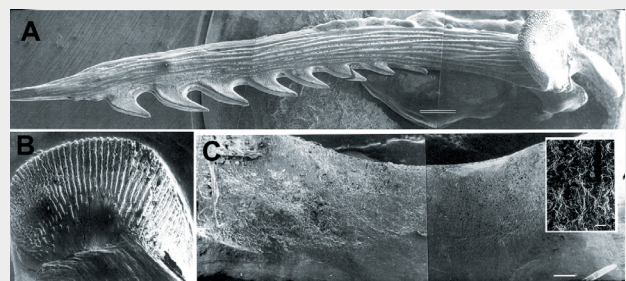
are thus highly variable and provide numerous examples of convergent evolution (independent evolution of similar traits in unrelated species), many of which center around the swim bladder.

Studies of fish sound production are going through a revolutionary period as they expand from a niche area with a small number of investigators. Reasons for this shift include the development of long-term recorders enabling the passive acoustic monitoring (PAM) (Mann, 2012) of species that cannot be easily observed, such as the monitoring of populations of Atlantic cod (*Gadus morhua*) and haddock (see dosits.org/Haddock) to locate breeding sites (e.g., Hawkins and Picculin, 2019). Using PAM thus provides ecological information on underwater habitats that would otherwise be difficult to obtain, e.g., nocturnal-burying or deep-living species that are hard to sample and cryptic species not resolved taxonomically (Parmentier et al., 2021b). Indeed, as early as the mid-twentieth century, the famous oceanographer N. B. Marshall (1967) identified muscles that are likely used for sound production in fishes living at great depth (macrourids and ophidiids), although we have yet to record identified sounds from these species in deep water.

The Swim Bladder
The swim bladder is situated below the vertebral column and forms as a dorsal outgrowth of the digestive tract. A major function is buoyancy; the internal gas counteracts the density of the heavier bone and muscle (Pelster, 2021). Like the digestive tract, the swim bladder wall is made up of cell types grouped in external and internal layers. Wall structure has not been correlated with acoustic function across teleost fishes, let alone been the subject of much experimentation. Swim bladders are practically transparent in shads (relatives of herring) that can hear 180-kHz ultrasounds (Popper et al., 2004) and have a heavy white external covering implicated in sound production and hearing in many fishes (Fine et al., 2016).

The gas in the swim bladder provides an acoustic discontinuity (impedance mismatch) to water (Urick, 1975). The swim bladder has been treated as an underwater resonant bubble (Sand and Hawkins, 1973), a monopole that radiates sound equally in all directions (Harris, 1964). Moreover, the resonant frequency of a bubble decreases with size and increases with depth (Minnaert, 1933).

Historically, swim bladders were believed to aid hearing by translating vibrations resulting from incident sound to the ears in all species. Although on a continuum, fishes vary from auditory specialists to generalists (Popper et al., 2021). In specialists, swim bladders are close to or connected to the ears and support sensitive thresholds and hearing to several kilohertz. Specialist connections vary from bony Weberian ossicles that function something like the mammalian middle ear bones in otophysans (e.g., minnows, goldfish, and catfish) (Braun and Grande, 2008) to diverticula, or tubular connections, in other families (Parmentier et al., 2011). By contrast, generalists have no connections between the swim bladder and the ear, have less sensitive hearing and typically hear sounds below 1 kHz. Indeed, experiments in which the swim bladder was deflated in generalists (e.g., oyster toadfish, blue gourami, a goby) do not change the shape or sensitivity of audiogram threshold curves (Yan et al., 2000), indicating that generalist swim bladders are unlikely to enhance hearing.

Significantly, connections between the swim bladder and ear can vary even between closely related species. For example, some members of the taxonomic family Sciaenidae, a group of fishes often called drums or croakers because of the loud sounds they produce (e.g., see dosits.org/Silver-Perch), have long diverticula that terminate close to the ears (e.g., silver perch, weakfish). These species have lower thresholds and higher-frequency reception than the related spot and Atlantic croaker that have shorter diverticula (Horodysky et al., 2008).

Sonic Swim Bladder Exaptations
Except for hydrodynamic (swimming) sounds, sonic mechanisms appear to be exaptations, a term referring to a functional character that is co-opted for a new use (Gould and Verba, 1982). Structures with other functions have been repurposed into sonic organs while still maintaining their original functions (Parmentier and Fine, 2016). The first description of a sonic exaptation came from the clownfish (think Nemo from the movies; see dosits.org/Clownfish) with ligaments that cause jaw slams for prey capture (Parmentier et al., 2007). The slams originally produced sounds incidentally (Olivier et al., 2014), but they are now used during courtship and territorial defense (Parmentier et al., 2021b). Another example is the pharyngeal teeth at the back of the mouth

in many fishes, which are used to crush food and guide it into the esophagus. In haemulids, commonly known as grunts, these teeth have assumed an additional function and can generate a series of disturbance pulses apart from feeding (see dosits.org/Bluestriped-Grunt) (Bertucci et al., 2014).

Most sonic muscles are extrinsic; their morphological origin is on various bones (e.g., skull, vertebrae, ribs) and their insertion is on the swim bladder or structures that move the swim bladder. Based on forming from neck muscles in the oyster toadfish, extrinsic muscles likely gave rise to intrinsic ones found only on the walls of the swim bladder (Mok et al., 2011). The evolutionary process of attaching muscles to a swim bladder is unclear, with the possible exception of piranhas and related species (Melotte et al., 2016) in which intercostal horizontal muscles between ribs first modified their orientation and later hypertrophied. In the most derived species (e.g., those with the greatest divergence from ancestral species), the sound-producing system has two symmetric vertically oriented muscles coming from the vertebrae and connected by a tendon surrounding the ventral swim bladder. The muscles produce sounds by rapidly compressing the base of the anterior swim bladder (Melotte et al., 2016). Interestingly, sonic muscles in a related characiform group (tetras) maintained their primitive horizontal orientation and compressed the front of the swim bladder, illustrating that evolution likely occurred in stages (Borie et al., 2019).

In other species, the evolutionary stages are unknown, but developmental data support swim bladder muscles forming from head or body muscles. In toadfish, for example, the occipital-spinal nerves and sonic muscles migrate from the neck region and attach to the swim bladder embryologically, whereas in sciaenids with true spinal innervation, the sonic muscles grow down from the dorsal aponeurosis (a flat tendon overlying the swim bladder) during puberty. The two basic innervation patterns (occipital-spinal and true spinal) have evolved independently multiple times (Rice et al., 2022).

Swim Bladder Shape and Sound Production

One of the most remarkable things about sound production in fishes is the extraordinary diversity in swim bladder shape and associated structures that are found in various species. Indeed, this reflects the extraordinary number of different approaches evolution has taken. Different groups of fishes have independently found multiple ways to produce sounds.

Swim bladders are typically somewhat circular in cross section, taper toward the tail, and slope downward (**Figure 1**), effectively lowering a fish's center of buoyancy. Because the swim bladder is restricted to the body cavity, the downward slope may provide lift to the tail region and the taper greater postural stability than a sharp termination. Both nonsonic and some sonic species have swim bladders with this general shape.

Sciaenid bladders and some others often terminate in a pointed "duck tail" that should allow minimal sound radiation and reflect gas pressure forward during muscle contraction (**Figure 1**). Sound amplitude is determined by volume velocity, a product of surface area of the swim bladder and velocity (Bradbury and Vehrencamp, 1998). Thus, the small surface area and a thick tail will produce lower amplitudes than the larger swim bladder sides moved by sonic muscles. Because sonic muscles surround the mid-to-posterior swim bladder in many sciaenids, contraction will compress the bladder, increasing the internal pressure and causing expansion of the anterior region not covered by muscles.

Several sciaenids have caudally directed diverticula (Chao, 1986). For example, the small scale croaker with six long slender diverticula bilaterally (**Figure 1**) produces sounds with missing or attenuated lower frequencies, suggesting that the diverticula are acting as Helmholtz absorbers (e.g., high-pass filters; Mok et al., 2020). Black drum (Locascio and Mann, 2011) and some doradid catfish swim bladders have many short blind diverticula that increase the surface area and maybe sound amplitude (Birindelli et al., 2009).

Modified swim bladder shapes may be related to sound production, audition, and/or gas secretion. Notably, the swim bladder of the oyster toadfish has a heart "cardiform" shape (**Figure 4**) with anterior protrusions on either side of the midline. Two hypotheses could explain this shape. The protrusions could direct sound detected by the swim bladder toward the ears, enhancing hearing. Alternatively, due to the absence of sonic muscles on the

Figure 4. *Dissected swim bladder of the oyster toadfish* Opsanus tau. *Anterior (front) part is on* **top.** *A: dorsal view.* **B:** *ventral view.* **C:** *ventral view with swim bladder floor removed, illustrating the heavy column between the two anterior projections (**arrowheads**). Scale bar, 1 cm. From Barimo and Fine (1998), used with permission.*

1998). Additional sonic fishes with somewhat similarly shaped bladders likely impart a directional sound field (Ladich and Fine, 2006).

There are also swim bladders with variable heart or cardiform shapes in catfishes (Birindelli et al., 2009), species not related to the aforementioned oyster toadfish. These may support auditory function in concert with the Weberian ossicles that transfer swim bladder vibration to the ears (Chardon et al. 2003). Catfish swim bladders have two internal septa; a longitudinal septum runs from the back of the bladder forward and splits into a horizontal septum, forming a "T." The top of the T extends laterally from the midline without reaching the side walls, leaving lateral channels. Because the middle part of the anterior bladder (forward of the horizontal septum) is stiffer than lateral areas near the attachment of the first ossicle (Ali and Fine, unpublished data), the lateral channels are interpreted as an adaptation to aid hearing by amplifying vibrations near the ossicle attachment point.

The oyster toadfish and some other species including some cichlids have a horizontal septum with a small hole surrounded by a sphincter that separates the anterior gas-secreting and posterior gas-absorbing regions (Tracy, 1911). Likely, this septum functions exclusively in gas regulation with no acoustic role. It therefore appears that internal swim bladder structures can have various acoustic and nonacoustic functions.

Many otophysans and members of unrelated families have bladders with one or more chambers (Melotte et al., 2016). The anterior chamber is likely an acoustic specialization. Related channel and blue (*Ictalurus punctatus* and *Ictalurus furcatus*, respectively) catfish have one- and two-chambered bladders, respectively (Ghahramani et al., 2014), illustrating swim bladder plasticity. Again, there is an overlap in the occurrence of hearing and sound production. In species without sonic muscles, the anterior chamber would be strictly auditory in function while still aiding buoyancy.

Various ophidiiform fishes (cusk-eels and carapids) have little external variation but have numerous sexually dimorphic swim bladder adaptations that may help them find mates in the dark using sound (**Figure 5**). These include a fenestra (a thin pliable strip with no outer layer) that allows sonic muscles to pull the anterior part of the swim

anterior bladder and the thick column connecting the two protrusions (**Figure 4**), sound could be directed away from the ears so that they are not overstimulated when the fish calls.

In testing these hypotheses, experiments in which the swim bladder was deflated did not show a change in auditory thresholds, failing to support the first hypothesis (Yan et al., 2000). By contrast, field recordings indicate that the sound amplitude is several decibels higher behind that of a toadfish as predicted by swim bladder shape and sonic muscle investment (Barimo and Fine,

Figure 5. *Top: left lateral view of the sound-producing apparatus of a male cusk-eel* Ophidion rochei *(Roche's snake blenny), showing the skeleton and swim bladder.* **Bottom:** *same structures with sonic muscles attached.* **Numbers I to V:** *first five vertebrae. The neural rocker is the neural arch over the first vertebra that rotates forward when the dorsal sonic muscle contracts. This movement causes backward movement of the winglike process (a modified rib on the first vertebra) and counterclockwise movement of the rocker bone at the front of the swim bladder. In this way, the dorsal sonic muscle contraction allows arming the sound-producing apparatus. During sustained contraction of the dorsal sonic muscle, rhythmic contraction cycles of the ventral sonic muscle causes back and forth movements of the rocker bone and the corresponding production of pulses. Figure by Eric Parmentier, © 2022, all rights reserved.*

bladder forward without moving the posterior part (Fine et al., 2007; Parmentier et al., 2010). Other families also have a fenestra (Parmentier et al., 2016), indicating that this adaptation evolved multiple times. Additional male ophidiform adaptations that occur in some species include a bean-shaped "rocker bone" formed by modification and sclerification of the anterior swim bladder wall (Parmentier et al., 2008), swim bladder tubercles (Parmentier et al., 2018), and a posterior round membrane within a raised "donut" that appears to be a pressure-release system (Kever et al., 2014; Ali et al., 2016). Some also have a pivoting neural arch and winglike process on the first vertebra, which along with antagonistic muscle pairs move the swim bladder in opposite directions (Fine et al., 2007; Kever et al., 2014). Muscle antagonists are unusual in sonic fishes that rely primarily on internal gas pressure to restore bladder shape during muscle relaxation.

Swim Bladders and the Underwater Resonant Bubble

By the 1960s, the notion that fish swim bladders acted as a pulsating resonant bubble was so entrenched in fish bioacoustics (e.g., Harris, 1964; Weston, 1967; Sand and Hawkins, 1973) that questioning the resonance model caused papers to be rejected for publication. Indeed, Tavolga (1971) supported resonance despite providing contrary evidence and suggested that the resonant frequency is close to the frequency produced by sonic muscles.

In the 1990s, we started to question the resonant bubble model as applied to sound production and hearing in the oyster toadfish. The bubble model predicts the swim bladder is an underdamped omnidirectional sound source that would emphasize a narrow frequency response based on size and depth. Findings of rapid damping had been attributed to swim bladder contact with surrounding fish tissue. Early work ignored the swim bladder wall, which is composed of collagen and elastin fibers (Morris and Albright, 1975). These fibers are woven in various directions and the wall contains about 80% water (Fine et al., 2016), therefore supporting the idea that viscous damping of the internal gas bubble inhibits the expression of resonance (Fine et al., 2016).

Our arguments and findings counter the traditional model and suggest a forced rather than a resonant response in toadfish and other species. These include the following observations.

(1) Most fish sounds are pulses with wide frequency spectra (Parmentier and Fine, 2016) rather than being concentrated at the swim bladder's resonant frequency. Tonal harmonic sounds such as the oyster toadfish boat whistle are caused by continuous contractions with a fundamental frequency that matches the muscle contraction rate (Fine et al., 2001). In fact, individual field recordings of male boat whistle choruses often have fundamental frequencies that vary by as little as 10 Hz, even though males are of different sizes (Fine, 1978). Therefore, muscle contraction rate rather than bladder size dictates fundamental frequency (Parmentier and Fine, 2016).

(2) The boat whistle fundamental frequency increases with temperature when muscles contract faster in toadfish (Fine, 1978) and other species (Ladich, 2018), although the resonant frequency of the internal bubble would be unaffected (Sprague et al., 2022).

(3) Toadfish sonic muscles are among the fastest vertebrate muscles and can be stimulated at 500 Hz without tetanizing (Fine et al., 2001). A resonant structure, however, does not require rapid excitation to set it into vibration (ding a crystal bowl).

(4) Toadfish calls terminate rapidly after muscle contractions stop (Fine et al., 2001), whereas a resonant structure would continue to ring, compromising the temporal nature of fish communication (Winn, 1964; Millot et al., 2011). Notably, toadfish can produce a grunt (acoustic tagging) within 41 ms in response to sounds from nearby toadfish or snapping shrimp (Thorson and Fine, 2002), suggesting precise temporal resolution that would be compromised by resonance.

(5) Oscillating bubbles are monopoles that produce an omnidirectional sound field (Harris, 1964). Yet the curved sonic muscles push the sides of the bladder inward, increasing the internal pressure that pushes the bottom outward (a quadrupole motion) (Fine et al., 2001). The sound radiation pattern in the wild is not uniform and follows the morphology of the heart-shaped swim bladder (Barimo and Fine, 1998).

(6) Deflation of the toadfish swim bladder does not affect the auditory threshold curve, although the swim bladder is several centimeters behind the ears (Yan et al., 2000).

Although a larger bubble will have a lower resonant frequency, low-frequency calls can occur for other reasons (Parmentier and Fine, 2016). Sonic muscles and swim bladders grow with fish size (Fine et al., 1990), and longer muscles take longer to contract, driving lower frequency sounds in weakfish (Connaughton et al., 2000) and other sciaenids (Tellechea et al., 2010). In a carapid fish, slow sonic muscles pull the anterior swim bladder forward, stretching the swim bladder fenestra that snaps back when a catch is released (Parmentier et al., 2006). The snap excites an overlying bony plate that, in turn, vibrates the swim bladder. The peak frequency of their sounds is close to twice that predicted by the bubble equation and is likely caused by tuning of the swim bladder plate. Finally, a cusk-eel living down to 5,000 m has short sonic muscles that terminate in a long tendon occupying 70% of the distance between the head and the swim bladder (Fine et al., 2018). Tendons require fewer calories than muscles and likely drive swim bladder sounds after the muscle contraction ends.

Clearly, there is support for both the forced response in sound production and hearing and the resonance model in sound fields. With depth, a bubble's resonant frequency will increase due to a higher hydrostatic pressure (Minnaert, 1933) and the gas will dominate so that the bladder wall will contribute less to the overall stiffness (Sprague et al., 2022). It therefore appears that the gas bubble resonance inside the bladder becomes increasingly important with depth, and both models likely apply, albeit under different environmental conditions.

Conclusion

Analogous to the syrinx in birds and larynx in mammals, the swim bladder is a major sound-producing organ in fishes. It also functions in hearing and is responsible for most returns from sonar. Historically, the swim bladder has been modeled as an underwater resonant bubble, whereas we provide evidence for a forced-response model in which viscous damping of the bladder wall inhibits resonance of the internal bubble. Further experimental work on the shape, wall structure, and mechanical behavior of the swim bladder is required.

References

Ali, H., Mok, H. K., and Fine, M. L. (2016). Development and sexual dimorphism of the sonic system in deep sea neobythitine fishes: The upper continental slope. *Deep-Sea Research Part I: Oceanographic Research Papers* 115, 293-308.

Barimo, J. F., and Fine, M. L. (1998). Relationship of swim-bladder shape to the directionality pattern of underwater sound in the oyster toadfish. *Canadian Journal of Zoology* 76, 134-143.

Bertucci, F., Ruppé, L., Van Wassenbergh, S., Compère, P., and Parmentier, E. (2014). New insights into the role of the pharyngeal jaw apparatus in the sound-producing mechanism of *Haemulon flavolineatum* (Haemulidae). *Journal of Experimental Biology* 217, 3862-3869.

Birindelli, J. L., Sousa, L. M., and Sabaj Perez, M. H. (2009). Morphology of the gas bladder in thorny catfishes (Siluriformes: Doradidae). *Proceedings of the Academy of Natural Sciences of Philadelphia* 158, 261-296.

Borie, A., Hungria, D. B., Ali, H., Doria, C. R., Fine, M. L., and Travassos, P. E., (2019). Disturbance calls of five migratory Characiformes species and advertisement choruses in Amazon spawning sites. *Journal of Fish Biology* 95, 820-832.

Bradbury, J. W., and Vehrencamp, S. L. (1998). *Principles of Animal Communication*. Sinauer Associates, Sunderland, MA.

Braun, C. B., and Grande, T. (2008). Evolution of peripheral mechanisms for the enhancement of sound reception. In Webb. J. F., Popper, A. N., and Fay, R. R. (Eds.), *Fish Bioacoustics*. Springer, New York, NY, pp. 99-144.

Chao, N. L. (1986). A synopsis on zoogeography of Sciaenidae. In Uyeano, T., Arai, R., Taniuchi, T., and Matsuura, K. (Eds.), *Indo-Pacific Fish Biology. Proceedings of the Second International Conference on Indo-Pacific Fishes*, Ichthyological Society of Japan, Tokyo, July 28 to August 3, 1985, pp. 570-589.

Chardon, M., Parmentier, E., and Vanderwalle, P. (2003). Morphology, development and evolution of the Weberian apparatus in catfish. In Arratia, G., Kapoor, B. G., Chardon, M., and Diogo, R. (Eds.), *Catfishes*, vol. 1. Science Publishers, Inc., Enfield, NH, pp. 71-120.

Connaughton, M. A., Taylor, M. H., and Fine, M. L. (2000). Effects of fish size and temperature on weakfish disturbance calls: Implications for the mechanism of sound generation. *Journal of Experimental Biology* 203, 1503-1512.

Fine, M. L. (1978). Seasonal and geographic variation of the mating call of the oyster toadfish *Opsanus tau. Oecologia* 36, 45-57.

Fine, M. L., and Parmentier, E. (2015). Mechanisms of fish sound production. In Ladich, F. (Ed.), *Sound Communication in Fishes*. Springer, Vienna, Austria, pp. 77-126.

Fine. M. L., Ali, H. A., Nguyen, T. K., Mok, H. K., and Parmentier, E. (2018). Development and sexual dimorphism of the sonic system in three deep-sea neobythitine fishes and comparisons between upper mid and lower continental slope. *Deep-Sea Research Part I: Oceanographic Research Papers* 131, 41-53.

Fine, M. L., Burns N. M., and Harris, T. M. (1990). Ontogeny and sexual dimorphism of the sonic muscle in the oyster toadfish. *Canadian Journal of Zoology* 68, 1374-1381.

Fine, M. L., Friel, J. P., McElroy, D., King, C. B., Loesser, K. E., and Newton, S. (1997). Pectoral spine locking and sound production in the channel catfish *Ictalurus punctatus. Copeia* 1997, 777-790.

Fine, M. L., King, T. L., Ali, H., Sidker, N., and Cameron, T. M. (2016). Wall structure and material properties cause viscous damping of swimbladder sounds in the oyster toadfish *Opsanus tau. Proceedings of the Royal Society of London B: Biological Sciences* 283, 1-9.

Fine, M. L., Lin, H., Nguyen, B. B., Rountree, R. A., Cameron, T. M., and Parmentier, E. (2007). Functional morphology of the sonic apparatus in the fawn cusk-eel *Lepophidium profundorum* (Gill, 1863). *Journal of Morphology* 268, 953-966.

Fine, M. L., Malloy, K. L., King, C. B., Mitchell, S. L., and Cameron, T. M. (2001). Movement and sound generation by the toadfish swimbladder. *Journal of Comparative Physiology A* 187, 371-379.

Ghahramani, Z. N., Mohajer, Y. J., and Fine, M. L. (2014). Developmental variation in sound production in water and air in the blue catfish *Ictalurus furcatus. Journal of Experimental Biology* 217, 4244-4251.

Gould, S. J., and Vrba, E. S. (1982). Exaptation — A missing term in the science of form. *Paleobiology* 8, 4-15.

Harris, G. G. (1964). Considerations on the physics of sound production by fishes. In Tavolga, W. N. (Ed.), *Marine Bio-Acoustics*. Pergamon Press, New York, NY, pp. 233-247.

Hawkins, A. D., and Picciulin, M. (2019). The importance of underwater sounds to gadoid fishes. *The Journal of the Acoustical Society of America* 146(5), 3536-3551. https://doi.org/10.1121/1.5134683.

Horodysky, A. Z., Brill, R. W., Fine, M. L., Musick, J. A., and Latour, R. J. (2008). Acoustic pressure and particle motion thresholds in six sciaenid fishes. *Journal of Experimental Biology* 211, 1504-1511.

Kever, L., Boyle, K. S., Bolen, G., Dragicevic, B., Dulčić, J., and Parmentier, E. (2014). Modifications in call characteristics and sonic apparatus morphology during puberty in *Ophidion rochei* (Actinopterygii: Ophidiidae). *Journal of Morphology* 275, 650-660.

Ladich, F. (2018). Acoustic communication in fishes: Temperature plays a role. *Fish and Fisheries* 19(4), 598-612.

Ladich, F., and Fine, M. L. (2006). Sound-generating mechanisms in fishes: A unique diversity in vertebrates. In Ladich, F., Colin, S. P., Moller, P., and Kapoor, B. G. (Eds.), *Communication in Fishes.* Science Publishers, Inc., Enfield, NH, pp. 3-43.

Locascio, J. V., and Mann, D. A. (2011). Localization and source level estimates of black drum (*Pogonias cromis*) calls. *The Journal of the Acoustical Society of America* 130, 1868-1879.

Looby, A., Cox, K., Bravo, S., Rountree, R., Juanes, F., Reynolds, L. K. and Martin, C. W., (2022). A quantitative inventory of global soniferous fish diversity. *Reviews in Fish Biology and Fisheries* 32, 581–595. https://doi.org/10.1007/s11160-022-09702-1.

Mann, D. A. (2012). Remote sensing of fish sounds using passive acoustic monitoring. *Acoustics Today* 8(3), 8-15.

Marshall, N. B. (1967). Sound-producing mechanisms and the biology of deep-sea fishes. In Tavolga, W. N. (Ed.), *Marine Bio-Acoustics,* vol. 2. Pergamon Press, New York, NY, pp. 123-133.

Melotte, G., Virourous, R., Michel, C., and Parmentier, E. (2016). Interspecific variation of warning calls in piranhas: A comparative analysis. *Science Reports* 6, 36127.

Millot, S., Vandewalle, P., and Parmentier, E. (2011). Sound production in red-bellied piranhas (*Pygocentrus natteri*, Kner): An acoustical, behavioural and morphofunctional study. *Journal of Experimental Biology* 214, 3613-3618.

Minnaert, F. (1933). On musical air bubbles and the sounds of running water. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics* 16, 235-248.

Mitchell, S., Poland, J., and Fine, M. L. (2008). Does muscle fatigue limit advertisement calling in the oyster toadfish *Opsanus tau*? *Animal Behaviour* 76, 1011-1016.

Mohajer, Y., Ghahramani, Z., and Fine, M. L. (2015). Pectoral sound generation in the blue catfish *Ictalurus furcatus. Journal of Comparative Physiology A* 201, 305-315.

Mok, H. K., Parmentier, E., Chiu, K. H., Tsai, K. E., Chiu, P. H., and Fine, M. L. (2011). An intermediate in the evolution of superfast sonic muscles. *Frontiers in Zoology* 8, 1-8.

Mok, H. K., Wu, S. C., Sirisuary, S., and Fine, M. L. (2020). A sciaenid swim bladder with long skinny fingers produces sound with an unusual frequency spectrum. *Science Reports* 10, 18619.

Morris, S. M., and Albright, J. T. (1975). The ultrastructure of the swimbladder of the toadfish, *Opsanus tau* L. *Cell and Tissue Research* 164, 85-104.

Olivier, D., Frédéric, B., Zarco, M. S., Balart, E. F., and Parmentier, E. (2014). The cerato-mandibular ligament: A key functional trait for grazing in damselfish (Pomacentridae). *Frontiers in Zoology* 11, Article 63.

Parmentier, E., and Fine, M. L. (2016). Fish sound production: Insights. In Suthers, R. A., Fitch, W. T., Fay. R, R., and Popper, A. N. (Eds.), *Vertebrate Sound Production and Acoustic Communication.* Springer, Cham, Switzerland, pp. 19-49.

Parmentier, E., Bahrib, M. A., Plenevaux A., Fine, M. L., and Estradad, J. M. (2018). Sound production and sonic apparatus in deep-living cusk-eels (*Genypterus chilensis* and *Genypterus maculatus*). *Deep-Sea Research Part I: Oceanographic Research Papers* 141, 83-92.

Parmentier, E., Bertucci, F., Bolgan, M., and Lecchini, D. (2021a). How many fish could be vocal? An estimation from a coral reef (Moorea Island). *Belgian Journal of Zoology* 151, 1-29.

Parmentier, E., Bouillac, G., Dragicevic, B., Dulcic J, and Fine, M. (2010). Call properties and morphology of the sound producing-organ in *Ophidion rochei* (Ophidiidae). *Journal of Experimental Biology* 213, 3230-3236.

Parmentier, E., Colleye, O., Fine, M. L., Frédérich, B., Vanderwalle, P., and Herrel, A. (2007). Sound production in the clownfish *Amphirprion clarkii*. *Science* 316, 1006.

Parmentier, E., Compère, P., Casadevall, M., Fontenelle, N., Cloots, R., and Henrist, C. (2008). The rocker bone: A new kind of mineralised tissue? *Cell and Tissue Research* 334, 67-79.

Parmentier, E., Fine, M. L, and Mok, H. K. (2016). Sound production by a recoiling system in the Pempheridae and Terapontidae. *Journal of Morphology* 277, 717-724.

Parmentier., E., Lagardère, J. P., Braquegnier, J. B., Vandewalle, P., and Fine, M. L. (2006). Sound production mechanism in carapid fish: First example with a slow sonic muscle. *Journal of Experimental Biology* 209, 2952-2960.

Parmentier, E., Mann, K. A., and Mann, D. A. (2011). Hearing and morphological specializations of the mojarra (*Eucinostomus argenteus*). *Journal of Experimental Biology* 214, 2697-2701.

Parmentier, E., Scalbert, R., Raick, X., Gache, C., Frédérich, B., Bertucci, F., and Lecchini, D. (2021b). First use of acoustic calls to distinguish cryptic members of a fish species complex. *Zoological Journal of the Linnean Society* 195(3), 964-975.

Pelster, B. (2021) Using the swimbladder as a respiratory organ and/or a buoyancy structure—Benefits and consequences. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology* 335, 831-842.

Popper, A. N., Hawkins, A. D., and Sisneros, J. A. (2021). Fish hearing "specialization" — A re-evaluation. *Hearing Research* 108393. https://doi.org/10.1016/j.heares.2021.108393.

Popper, A. N., Plachta, D. T., Mann, D. A., and Higgs, D. (2004). Response of clupeid fish to ultrasound: A review. *ICES Journal of Marine Science* 61, 1057-1061.

Rice, A. N., Farina S. C., Makowski, A. J., Kaatz, I. M., Lobel, P. S., Bemis, W. E., and Bass, A. H. (2022). Evolutionary patterns in sound production across fishes. *Ichthyology and Herpetology* 110, 1-12.

Sand, O., and Hawkins, A. D. (1973). Acoustic properties of the cod swim bladder. *Journal of Experimental Biology* 58, 797-820.

Sprague, M. W., Fine, M. L.,and Cameron, T. M. (2022). An investigation of bubble resonance and its implications for sound production by deepwater fishes. *PLoS ONE* 17(7), e0267338.

Tavolga. W. N. (1971). Sound production and detection. In Hoar, W. S., and Randall, D. J. (Eds.), *Fish Physiology*. Academic Press, New York, NY, pp. 135-205.

Tellechea, J. S., Martinez, C., Fine, M. L., and Norbis, W. (2010). Sound production in the whitemouth croaker and relationship between fish size and call characteristics. *Environmental Biology of Fishes* 89, 163-172.

Thorson, R. F., and Fine, M. L. (2002). Acoustic competition in the gulf toadfish *Opsanus beta*: Acoustic tagging. *The Journal of the Acoustical Society of America* 111, 2302-2307.

Tracy, H. C. (1911). The morphology of the swim bladder in teleosts. *Annals of the New York Academy of Sciences* 38, 600-606, 638-649.

Urick, R. J. (1975). *Principles of Underwater Sound.* McGraw-Hill, New York, NY.

Weston, D. E. (1967). Sound propagation in the presence of bladder fish. In Albers, V. M. (Ed.), *Underwater Acoustics,* vol. 2. Plenum Press, New York, NY, pp. 55-88.

Winn, H. E. (1964). The biological significance of fish sounds. In Tavolga, W. N. (Ed.), *Marine Bio-Acoustics*. Pergamon Press, New York, NY, pp. 213-231.

Yan, H. Y., Fine, M. L., Horn, N. S., and Colon, W. E. (2000). Variability in the role of the gasbladder in fish audition. *Journal of Comparative Physiology A* 186, 435–445.

About the Authors

Michael L. Fine mfine@vcu.edu

Department of Biology
Virginia Commonwealth University
Richmond, Virginia 23062, USA

Michael L. Fine is a professor emeritus in the Department of Biology, Virginia Commonwealth University (Richmond). Dr. Fine received his BS in zoology from the University of Maryland, College Park (College Park), his MA in marine science from the College of William and Mary (Williamsburg, Virginia), and his PhD in oceanography from the University of Rhode Island (Kingston). He did a postdoc in auditory neurophysiology at Cornell University (Ithaca, New York), recording from the toadfish VIIIth nerve. At Virginia Commonwealth University, he worked on anatomy, sonic muscle physiology, and sound production in the oyster and gulf toadfish and later in other fish species.

Eric Parmentier
E.Parmentier@uliege.be

Department of Biology, Environment and Ecology
University of Liège
B-4000 Liège, Belgium

Eric Parmentier is a full professor and the principal investigator in the Laboratory of Functional and Evolutionary Morphology, University of Liège (Liège, Belgium). Prof. Parmentier received his PhD in science from the University of Liège. He did a postdoc in acoustic communication at Virginia Commonwealth University (Richmond) and on hearing abilities in fishes at the University of South Florida (Tampa). He is working on acoustic communication, the anatomy of sound-producing muscles in different teleost species, and passive acoustic monitoring.

Speech, Rhythm, and the Brain

Steven Greenberg

Introduction

Think of "rhythm" and what most likely comes to mind are music and dance. We intuitively "know" what good rhythm is, especially when it comes to entertainment. Indeed, rhythm is vital for the expression of emotion in the arts. But what often goes unappreciated is that rhythm also plays an important role in various forms of acoustic signaling, including spoken language and non-human communication (Kotz et al., 2018).

What is it about rhythm that accounts for its prevalence across the animal kingdom (Ravignani et al., 2019)? And why is it especially important for human communication?

Although definitive answers lie outside the scope of the present discussion, several of these issues are examined here through the lens of speech acoustics, perception, and neuroscience. It is argued that rhythm lies at the very heart of what makes humans especially adept at communication, binding sensory signals across modalities and linking such input with internal, often rhythmic, neural activity in the brain.

What Rhythm Is

For illustrative purposes, I begin our survey by examining rhythm from a *musical* perspective, distinguishing between two "flavors" of rhythm, the "cognitive" and the "physical." Cognitive rhythm is associated with musical elements like notes, accents, beats, measures, and phrases. Physical rhythmic elements are intensity, duration, interval, and modulation.

Musicologists have traditionally viewed rhythm as operating on a sequence of perceptual elements: "Rhythm may be defined as the way in which one or more unaccented beats are grouped in relation to an accented one… A rhythmic group can be apprehended only when its elements are distinguished from one another, rhythm… always involves an interrelationship between a single, accented (strong) beat and either one or two unaccented (weak) beats" (Cooper and Meyer, 1960, p. 6).

Within this cognitive framework, rhythm is deemed a *relational* property, one that governs how elements (e.g., musical notes, measures, phrases) interact with each other perceptually and cognitively. Such operations likely involve widespread communication across a constellation of brain centers associated with the senses, memory, and movement.

But rhythm doesn't function simply as a relational quality: "…rhythm is the one indispensable element of all music. Rhythm can exist without melody, as in the drumbeats of so-called primitive music, but *melody cannot exist without rhythm*. In music that has both harmony and melody, the rhythmic structure cannot be separated from them" (emphasis added) (Crossley-Holland, 1998; 2002; 2020).

In other words, rhythm serves as a unifying, global function, integrating different musical elements into a perceptual experience greater than the sum of its constituent parts. Precisely how rhythm performs this cognitive "magic" is not well understood (Ding et al., 2017). One possibility is that certain key physical elements of musical and speech rhythm "trigger" endogenous synchronous activity, or neural "oscillations," associated with the encoding and retrieval of information pertinent to a variety of sensory and cognitive experiences discussed in **Speech Rhythms in the Brain**.

Speech Rhythm and Linguistic Representations

What pertains to music also applies to speech; however, the specifics differ. Talkers do not generally speak in musical notes or measures, although poetic rhythm can reenforce emotion or be used to conjure imagery and scenarios (see youtu.be/S0mwhkv9ves for an online discussion) (Obermeier et al., 2013).
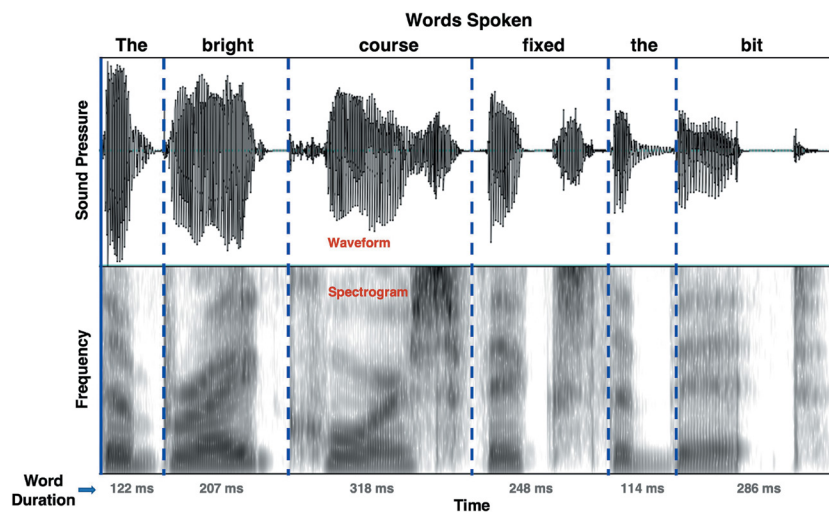
Figure 1. *The speech waveform (**top**) and associated spectrogram (**bottom**) for a sample speech signal. The words spoken are indicated above the waveform, which consists of both fast and slow modulations. The slower ones reflect syllabic and segmental rhythms. **Dotted vertical blue lines** separate adjoining words (which are also single syllables). Their durations are shown below the spectrogram.*

Here I examine how linguistic elements such as "phonetic segments," syllables, words, phrases, and sentences are impacted by the cognitive form of rhythm ("prosody") as well as by several physical attributes: modulation, phase, duration, frequency, and intensity.

Rhythm's physical form can be visualized via the acoustic signal's waveform (**Figure 1**). It contains both fast (i.e., higher frequency) and slow (i.e., very low frequency) sound pressure fluctuations. The fast modulations, the "temporal fine structure," are often associated with pitch and other tonal properties (Smith et al., 2002) but lie outside the scope of the present discussion.

Rhythm is reflected in the very slow modulations in the waveform, known as the "speech envelope," and in the motion of the speech articulators (Tucker and Wright, 2020), especially the opening and closing of the jaw, as well as the movement of the lips and tongue during speech production (Stevens, 1998). These parallel movements are the acoustic expression of speech rhythm. There is also a highly visible component, the so-called "speech-reading" cues associated with the articulatory movements that interact with certain elements of the acoustic signal to produce perceptual "objects" at the phonetic (van Wassenhove et al., 2007) and lexical (Winn, 2018) levels. The interaction

between the audio and visual speech signals, especially under challenging listening conditions, shields the speaker's message from the deleterious impact of background noise and other forms of acoustic interference (Assmann and Summerfield, 2004), something especially important for the hearing impaired.

The Dynamics of Rhythm

The motion of the articulators, especially the jaw, establishes the upper and lower bounds of the speech envelope's energy swings. These slow articulatory movements largely coincide with the linguistic element known as the "syllable." Although a syllable may contain just a single phonetic segment (e.g., "a") or as many as seven (e.g., "strengths"), most syllables contain just two or three (Greenberg, 1999). Although the average duration of a syllable is about 200 ms in American English (Greenberg, 1999) and 165 ms in Japanese (Arai and Greenberg, 1997), their length can vary from about 100 ms to about 330 ms. Such durational properties are important for the next discussion because they can also be expressed in terms of "modulation frequency," a key quantitative metric for representing speech rhythms across a range of temporal scales and is also important for speech intelligibility (the ability to decode and understand the words spoken in a phrase, sentence, or longer utterance).

In modulation-frequency units, syllables range between 3 Hz (for long-duration examples) and 10 Hz (for short-duration examples) (**Figure 2**). Syllables form the backbone of speech's modulation spectrum, a reflection of the articulatory dynamics associated with the opening and closing of the jaw during speaking, which modulates the amplitude of the acoustic signal. The intensity of a speech sound is closely related to the aperture of the oral cavity. More energy is released during the vocalic portion of the syllable when the opening is wide, whereas much less energy is released when the aperture is reduced during the production of (most) consonants. Hence, one can liken a syllable's waveform to an "energy arc" (Greenberg, 2006) where there are rises and falls in energy that closely follow the amplitude characteristics of the individual phonological constituents within a syllable. This is illustrated for the two-syllable word "seven" in a three-dimensional representation of the speech signal called a "spectro-temporal profile" (STeP; **Figure 3**). The STeP shows the energy dynamics of the speech signal using hundreds of instances of the same word that have been averaged to derive a composite representation (Greenberg et al., 2003).

Waveform modulations are also associated with a syllable's constituent phonetic segments (or "phones"), ranging in duration between ca. 50 ms (20 Hz) and ca. 150 ms (7 Hz). These faster undulations are nested within
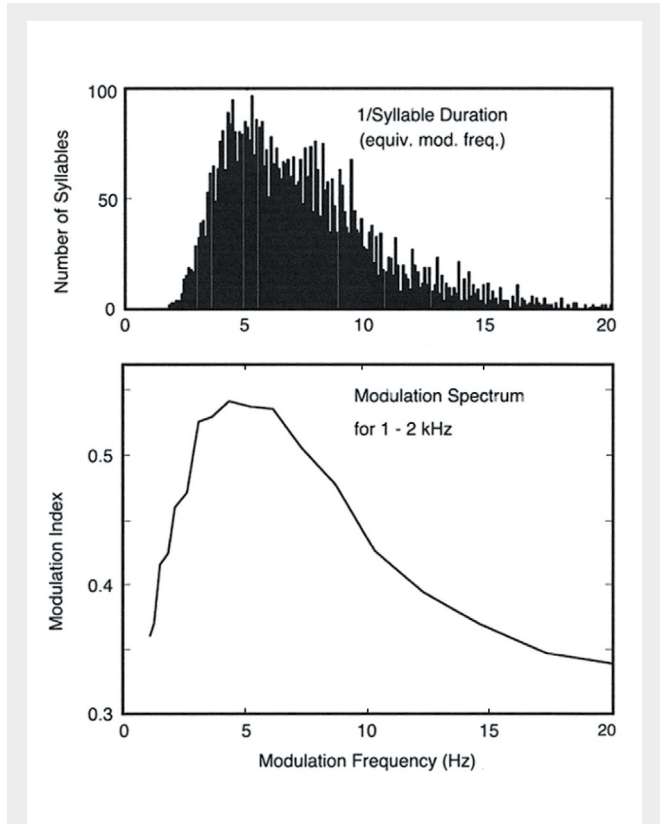
Figure 2. *The relationship between the distribution of syllable duration (transformed into equivalent modulation frequency [equiv. mod. freq.] units) (**top**) and the modulation spectrum of the same material (Japanese spontaneous speech) (**bottom**) as calculated for the octave region between 1 and 2 kHz.*
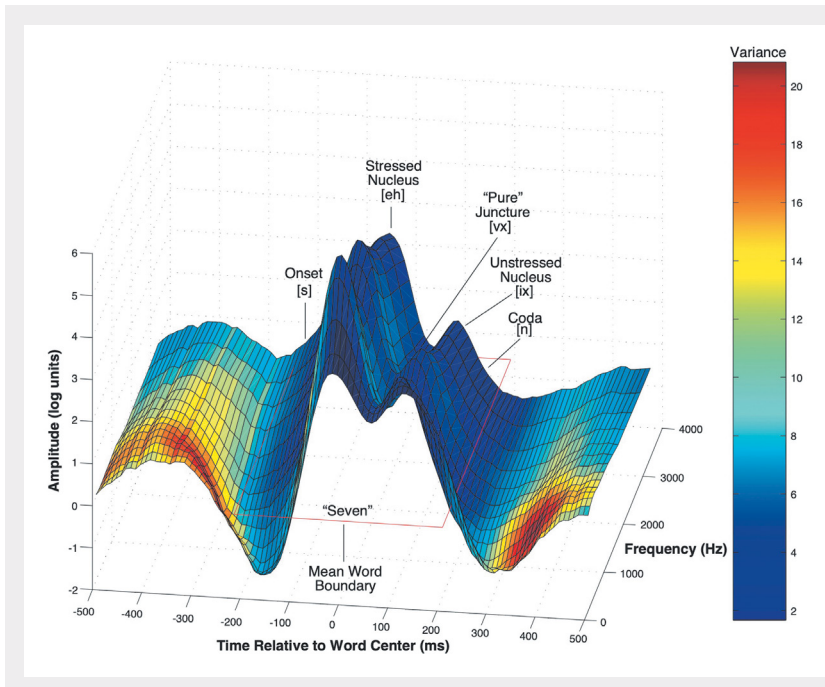
Figure 3. *A spectro-temporal profile (STeP) of the word "seven," a normalized averaging of hundreds of instances from the OGI Numbers corpus. The STeP shows the signal modulation patterns associated with the onset (s), nucleus, and coda (n) constituents of two syllables, the first stressed (eh) and the second unstressed (ix), to highlight the waveform dynamics of the spoken material. The pure juncture lies in the trough between the stressed and unstressed vocalic nuclei.*

syllabic modulations, imparting a phonetic detail required to achieve lexical clarity and semantic precision.

Fluctuations on a longer timescale than the syllable are often referred to as "prosodic," although there may be modulatory patterns within a syllable that are also of prosodic significance. These prosodic patterns are reflected in the modulation spectrum's lower limb (<3 Hz). Perceptually, these very low frequency modulations are instantiated in a syllable's *prominence* relative to neighboring syllables in a word, phrase, or sentence. These emphasized syllables are "accented" or "stressed" (Beckman, 1992). The intensity and duration of a syllable's vocalic core (known as the "nucleus") relative to nearby nuclei are the most important physical attributes of prominence (Silipo and Greenberg, 1999), although other physical properties play a role and have been incorporated into an automatic prosodic prominence labeling system, AutoSAL (Greenberg, 2005, Fig. 11).

It is not just the energy within an utterance that varies, but also its fundamental frequency (f_0; "pitch") contour. Such pitch variation may mark the transition from one grammatical phrase to another (tone and break indices [ToBI]; Silverman et al., 1992), helping the listener parse the speech signal for better comprehension. A computational version (AuToBI) uses pitch contour patterns as well as syllable duration and intensity to parse utterances (Rosenberg, 2010).

How important are speech rhythm and slow waveform modulations for intelligibility? As early as 1939, a Bell Labs engineer, Homer Dudley, recognized the importance of slow modulations for creating intelligible speech with his invention of the "vocoder" (Dudley, 1939). He distinguished between the fast-moving "carrier" (i.e., the spectro-temporal "fine structure") and the more slowly moving "modulator," making it clear that both are essential for creating intelligible speech (Bunnell, 2022). A vocoder consists of a series of band-pass filters, simulating the frequency analysis of the auditory system used to create a perceptual model of the speech signal that is more compact than the original. Modern-day applications of the vocoder are found in a variety of text-to-speech applications (Kawahara, 2015) and have been fine tuned to create much more natural sounding speech than Dudley's (1939) original version.

Dudley's (1939) insight received renewed interest in the 1970s when an automated system was developed for predicting intelligibility in acoustic environments like concert halls, theaters, and worship spaces (Houtgast and Steeneken, 1973). Key to the system's success was a method for quantifying the amount of energy in each frequency channel of the very slowly moving modulations in the speech signal. Houtgast and Steeneken dubbed their metric the "modulation spectrum" because it quantified the amount of energy in each frequency channel of modulation. In this context, "frequency channel" refers to a unit of time considerably longer (50 ms to 2 s) than the temporal units associated with tonal spectral audibility (50 µs to 20 ms) in human listeners. Houtgast and Steeneken noted that the contour of the modulation spectrum could be used to distinguish intelligible from unintelligible speech, especially in noisy and reverberant environments.

Why does the modulation spectrum's profile predict speech intelligibility so well? An intuitive explanation is that speech energy (i.e., acoustic "reflections") added back to the speech signal with a certain delay smooths the contours of the slow modulations in ways that degrade critical linguistic information within the syllable. The waveform modulations containing critical phonetic cues are no longer crisply defined, thereby compromising a listener's ability to extract sufficient phonetic detail to decode and interpret the speech signal. This intuition is consistent with a study by Drullman et al. (1994), who low-pass filtered the slow modulations using a procedure that "smeared" (i.e., "blurred") the boundaries between adjacent syllables, thereby distorting speech-relevant information in the modulation "packets." **Figure 4** shows how intelligibility declines as the *complex* modulation spectrum diminishes in amplitude (Greenberg and Arai, 2004).

The modulation spectra of intelligible speech material exhibit a peak between 4 and 8 Hz, the key range for syllabic information. There is also complex modulation energy between 8 and 16 Hz, the region most closely associated with phonetic segmental cues. The lower branch of the modulation spectrum (<4 Hz) is associated with highly prominent (i.e., accented) syllables. When the intelligibility decreases, so does the amount of energy in the modulation spectrum, especially in the critical 4- to 8-Hz region.

The importance of slow modulations for intelligibility was demonstrated in a different way by Shannon et al. (1995). In place of a conventional speech waveform with
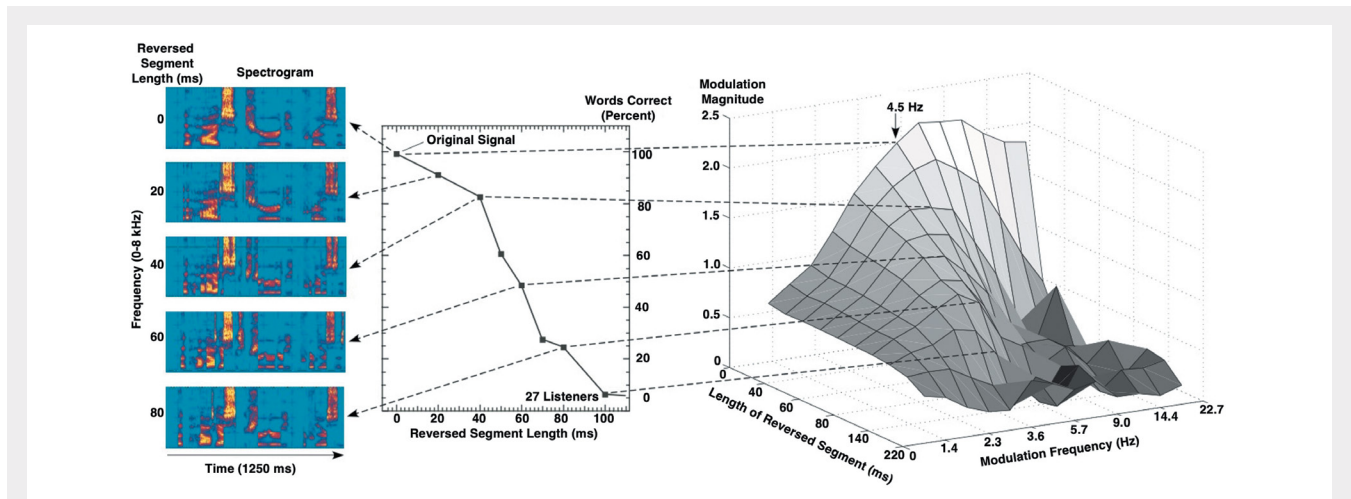
Figure 4. *The relationship between the complex magnitude of the modulation spectrum and speech intelligibility. The complex modulation spectrum integrates the magnitude and phase components into a single value. The sentence material's intelligibility was manipulated by locally time- reversing the speech signal over different segment lengths. As the reversed-segment duration increases beyond 40 ms, intelligibility declines precipitously, as does the magnitude of the complex modulation spectrum. The spectro-temporal properties also deteriorate appreciably under such conditions. Reprinted from Greenberg and Arai (2004).*

its harmonic (i.e., "voiced") structure, the carrier signal used was white noise. But the modulator of the original speech signal was retained, used to modulate the white-noise carrier in ways reminiscent of a coarse-grained spectrum analyzer. These slow modulations vary depending on whether they are derived from the low-, mid-, or high-frequency region of the speech signal's acoustic spectrum. Shannon et al. (1995) discovered that intelligible speech was only possible if the slow modulations were combined across different regions of the acoustic frequency spectrum. In other words, *a diverse set of slow modulators was required to preserve the linguistic information* contained in the original signal when transformed into a vocoded, noise-excited version.

To summarize, these pioneering studies demonstrated that low-frequency modulations in the speech waveform convey information critical to intelligibility. But these early demonstrations left unaddressed a variety of questions regarding how such information unlocks neurological pathways involved in speech comprehension and understanding.

It was at this point that my colleagues and I performed several studies to shed more light on *why* these slow modulations figure so importantly in speech perception. We asked five basic questions.

Q1: How much can the slow modulations be perturbed without impacting intelligibility?

Q2: Does the modulation spectrum vary across acoustic (and hence auditory) frequency?

Q3: How do the modulations across the acoustic frequency spectrum interact with each other?

Q4: Can the visual modality interact with the auditory speech signal to provide a measure of redundancy?

Q5: Can such perceptual data be linked to such linguistic "objects" as the phonetic segment, syllable, and word?

Our studies showed that

A1: The slow modulations can be perturbed only to a limited degree without seriously compromising intelligibility (Greenberg and Arai, 2004).

A2: The modulation spectrum does indeed vary appreciably across the acoustic frequency spectrum (Greenberg et al., 2003).

A3: Intelligibility is moderately sensitive to the *phase* (i.e., timing) of the slow modulations across acoustic frequency (Greenberg and Arai, 2004).

A4: Visual speech rhythmic patterns do interact with the acoustic signal but in an asymmetrical way. Intelligibility is far more tolerant of audiovisual asynchrony when the visual component leads the audio rather than vice versa (Grant and Greenberg, 2001).

A5: Different parts of the modulation spectrum are associated with distinct linguistic elements (Greenberg et al., 2003). Highly prominent (i.e., accented, stressed) syllables are associated with the lower limb (3-5 Hz) of the modulation spectrum, whereas less prominent syllables are associated with its upper limb (6-8 Hz). A syllable's prominence influences the phonetic realization of both consonant and vocalic segments (Greenberg, 2005).

Speech Rhythm in Broader Perspective

Prosody's power to connect with listeners is well-known to those engaged in entertainment, politics, or preaching. Much of this emotive force is an embodiment of specific properties of the speech signal, especially the emphasis placed on specific syllables (and words) and their timing relative to their less prominent counterparts. Such rhythms can help the listener navigate the speech stream to separate the semantic "wheat" from the "chaff."

Parsing and chunking speech are important for other reasons too. Speech is inherently ambiguous. Listen to a short snippet of, say, a single syllable or word and try to guess what the speaker is saying. This experiment was performed nearly 60 years ago by Pickett and Pollack (1963) and Pollack and Pickett (1964) who found that several words in succession (ca. 1 s) were required to reliably recognize the words in both read and conversational speech (**Figure 5**). This ambiguity of individual linguistic elements places a premium on predicting which sounds and words are likely to follow. Human listeners routinely do this, and recent brain-imaging data show that parts of the frontal cortex, especially the prefrontal region, are heavily involved (Park et al., 2015). This is where speech rhythm may play an especially important role because it serves as an organizing, active framework for critical cerebral centers to "latch on" to relevant neural activity in various parts of the brain (Schroeder et al., 2010).

However, speech rhythm entrails more than syllabic prominence and pitch contours. The emotional valence of what is said can be equally (if not more) important. The prosody of emotion has long been the subject of
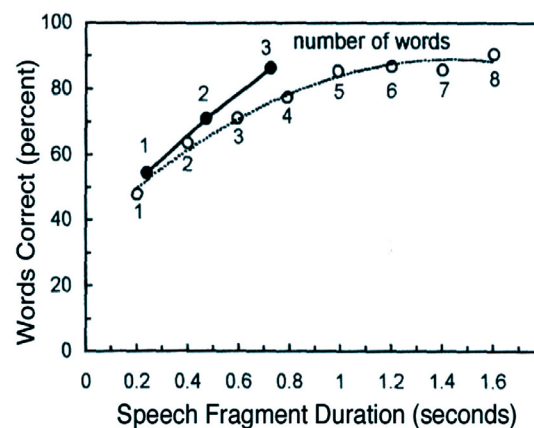
Figure 5. *Average identification score of words in fragments excised from read text (**closed circles**) and conversational speech (**open circles**) as a function of fragment duration. Based on data from Pickett and Pollack (1963) and Pollack and Pickett (1963) and adapted from Plomp (2002, p. 107).*

study and speculation (Scherer, 2003). Darwin (1873) was perhaps the first to suggest that emotion has deep roots in our phylogenetic history and that primitive elements of prosody can be found in vocalizations of certain nonhuman species (Ravignani et al., 2019).

From this global perspective, rhythm can serve as both a mediating and unifying force. It acts as a mediator between the lower level, physical and sensory tiers and the higher, cognitive levels associated with semantic and situational analysis and interpretation (Hawkins, 2014). Rhythm is a unifier in that it combines what might otherwise be just an assortment of unrelated acoustic elements (e.g., harmonics and other frequency components) and groups them together to create sensory objects capable of signaling words, phrases, and concepts (Elhilali, 2019).

Speech Rhythms in the Brain

How is speech processed by the brain? Are exogenous signal properties, such as syllabic modulations, linked to endogenous neural activity associated with linguistic functions like phonetic analysis, word recognition, and semantic interpretation (Zhang and Ding, 2017)? Can neurological investigations elucidate not only the pertinent brain mechanisms (Friederici, 2011) but also shed light on acoustic biocommunication in nonhuman

species (Ravignani et al., 2019)? Possibly, as there are acoustic properties shared across many species, and the way different parts of the brain communicate with each other also appears to be similar across much of the animal kingdom.

It has long been known that low-frequency rhythms can be recorded from the scalp of human subjects (Berger, 1929), although their significance remained unclear for many decades. Over the past 30 years, many different brain rhythms have been studied (Buzsaki, 2006). They range in frequency from the relatively fast, gamma-ϒ (ca. 30-80 Hz) to the very slow, delta-δ (0.5-3 Hz) and to points in between: theta-θ (3-8 Hz), alpha-α (8-10 Hz), and beta-β (10-20 Hz) (**Figure 6**).

These brain rhythms are not perfectly periodic but fluctuate around an average frequency, with energy spanning a range of spectral components. But the rhythms of speech also deviate from lockstep periodicity. When rhythm is studied, it is the central tendency rather than a metronomic pattern that forms the focus of analysis. Thus, it

Figure 6. *Hypothesized relationship between brain rhythms and speech processing over a range of timescales and neural oscillation frequencies. These neural oscillations reflect the synchronous activity of thousands (or millions) of neurons in the cerebral cortex and hippocampus responding to sensory stimulation (in this example, an acoustic speech signal). The different timescales of the oscillations are hypothesized to match the timescales of linguistic elements thought to be important for decoding and understanding the speech signal. Waveforms shown are solely for illustrative purposes.*

is not surprising that rhythms internal to the brain don't follow a strict timetable but rather largely reflect synchronous, endogenous communication among cerebral regions pertinent to the behavior at hand.

One of these endogenous rhythms, theta–θ, closely emulates the timing of syllables in spoken material (Poeppel and Assaneo, 2020) and may be "entrained" (i.e., extremely synchronized) to the signal's syllabic modulations. There are brain rhythms whose temporal properties are comparable to other linguistic elements, both shorter (phonetic segments; beta-β [50-100 ms, 10-20 Hz]) and longer (phrases; delta-δ [300-2,000 ms, 3-5 Hz]) than syllables (Etard and Reichenbach, 2019).

What is the significance of these neural oscillations? Are they merely tracking the signal's physical properties or do they reflect deeper processing germane to the analysis and interpretation of the linguistic message? Some studies indicate a marked entrainment to the waveform's syllabic rhythm on a cycle-by-cycle basis based on a variety of recording methods including electroencephalography (Fujii and Wan, 2014), magnetoencephalography (Poeppel and Assaneo, 2020), and electrocorticographic (Oganian and Chang, 2019).

However, there is evidence that at least some of these oscillations are linked to more profound processing, persisting well after cessation of the signal (van Bree et al., 2021). Several brain recording studies suggest that speech comprehension is most closely reflected in beta–β oscillations (Pefkou et al., 2017) and that a talker's speaking rate is faithfully reflected in theta–θ oscillations, which have also been linked to attentional processes (Fiebelkorn and Kastner, 2019) and is associated with parsing the speech into meaningful "chunks." Endogenous rhythms may also enhance the forecasting of phonetic segments, syllables, words, and phrasal structure in conversational speech. But how this linguistic "magic" is achieved is not currently well understood.

Rhythm in Developmental Perspective

Language, both spoken and written, requires time to acquire and master even by native speakers. It may not be until the age of 11 years that the child's linguistic prosody is fully formed (Polyanskaya and Ordin, 2015). Is it possible that the acquisition and mastery of a language depends on learning its rhythm? And if there is

some flaw in this skill might this deficit impair linguistic competence, at least for native speakers? There is some evidence that this is indeed the case, both in speech production (Fujii and Wan, 2014) and in reading (Leong and Goswami, 2014). Perhaps rhythm is a foundational property, one that holds the key to understanding language's neural bases.

The "Why" of Rhythm

Virtually all animals move, and such locomotion involves rhythmic motor activity, posing a challenge for sensory systems tasked with maintaining the illusion of stability for constantly changing stimuli. One way in which the brain can navigate such sensorimotor dynamics is through rhythmic patterns of neural activity (Lubinas et al., 2022) that submerge the intrinsic variability of sensory signals within nested hierarchies of cortical oscillations (Ghitza, 2011) that "translate" lower level features into more global, complex features of variable duration and cognitive complexity (Greenberg, 2011). Consistent with this perspective is a study that artificially distorted the rhythm of spoken sentences to disrupt intelligibility over the temporal range in which theta–θ oscillations are thought to operate (Ghitza and Greenberg, 2009). Perhaps the temporal patterning of spoken (and other forms of) communication evolved to "piggyback" on intrinsic rhythms of the brain (Kotz et al., 2018).

Rhythm's Future

Rhythm played a supporting role in the study of spoken language for most of the twentieth century, its importance only coming to the fore in the 1990s as perceptual and statistical studies highlighted rhythm's centrality for speech intelligibility and understanding. In recent years, this recognition has played a key role in integrating rhythm into speech synthesis technology to create more natural-sounding material (Bunnell, 2022) as well as incorporating rhythm into automatic speech-recognition models. Speech rhythm has also begun to be used in speech rehabilitation (Fujii and Wan, 2014), in foreign language instruction (Greenberg, 2018), and as an adjunct for teaching kids to read. And rhythm is now the center of attention for evolutionary studies of animal communication and its importance for the evolution of human language (Ravignani et al., 2019). The science of rhythm is in its infancy and is likely to provide further insights into language and other aspects of human behavior for years to come.

References

Arai, T., and Greenberg, S. (1997). The temporal properties of spoken Japanese are similar to those of English. *Proceedings of the Fifth European Conference on Speech Communication and Technology (Eurospeech 1997)*, Rhodes, Greece, September 22-25, 1997, pp. 1011-1014.

Assmann, P., and Summerfield, Q. (2004). The perception of speech under adverse conditions. In Greenberg, S., Ainsworth, W. A., Popper, A. N., and Fay, R. R. (Eds.), *Speech Processing in the Auditory System*. Springer, New York, NY, pp. 231-308.

Beckman, M. E. (1992). *Stress and Non-Stress Accent*. Fortis, Dordrecht, The Netherlands.

Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten* 87, 527-570. https://doi.org/10.1007/BF01797193.

Bunnell, H. T. (2022). Speech synthesis: Toward a "voice" for all. *Acoustics Today* 18(1), 14-22. https://doi.org/10.1121/AT.2022.18.1.14.

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford University Press, New York, NY.

Cooper, G., and Meyer, L. B. (1960). *The Rhythmic Structure of Music*. University of Chicago Press, Chicago, IL.

Crossley-Holland, P. (1998/2002/2020). Rhythm. *Encyclopedia Britannica*. Available at https://www.britannica.com/art/rhythm-music.

Darwin, C. (1871). *The Expression of Emotions in Man and Animals*, 4th ed. Oxford University Press, Oxford, UK.

Ding N., Patel, A. D., Chen, L., Butler, H., Luo, C., and Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience and Biobehavioral Reviews* 18, 181-187. https://doi.org/10.1016/j.neubiorev.2017.02.011.

Drullman, R., Festen, J. M., and Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *The Journal of the Acoustical Society of America* 95, 1053-1064. https://doi.org/10.1121/1.408467.

Dudley, H. (1939). Remaking speech. *The Journal of the Acoustical Society of America* 11, 169-177.

Elhilali, M. (2019). Modulation representations for speech and music. In Siedenburg, K., Saitis, C., McAdams, S., Popper, A. N., and Fay, R. R. (Eds.), *Timbre: Acoustics, Perception, and Cognition*, Springer, Cham, Switzerland, pp. 335-359. https://doi.org/10.1007/978-3-030-14832-4_12.

Etard, O., and Reichenbach, T. (2019). Neural speech tracking in the theta and in the delta frequency band differentially encode clarity and comprehension of speech in noise. *The Journal of Neuroscience* 39, 5750-5759. https://doi.org/10.1523/JNEUROSCI.1828-18.2019.

Fiebelkorn, I. C., and Kastner, S. (2019). A rhythmic theory of attention. *Trends in Cognitive Sciences* 23(2), 87-101. https://doi.org/10.1016/j.tics.2018.11.009.

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews* 91(4), 1357-1392. https://doi.org/10.1152/physrev.00006.2011.

Fujii, S., and Wan, C. Y. (2014). The role of rhythm in speech and language rehabilitation: The SEP hypothesis. *Frontiers of Human Neuroscience* 8, 777. https://doi.org/10.3389/fnhum.2014.00777.

Ghitza, O. (2011). Linking speech perception and neurophysiology: Speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in Psychology* 2, 130. https://doi.org/10.3389/fpsyg.2011.00130.

Ghitza, O., and Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: Intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica* 66, 113-126. https://doi.org/10.1159/000208934.

Grant, K.W., and Greenberg, S. (2001). Speech intelligibility derived from asynchronous processing of auditory-visual information. *Proceedings of the International Conference on Auditory-Visual Speech Processing,* Aalborg, Denmark, September 7-9, 2001, pp. 132-137.

Greenberg, S. (1999). Speaking in shorthand — A syllable-centric perspective for understanding pronunciation variation. *Speech Communication* 29, 159-176.

Greenberg, S. (2005). From here to utility — Melding phonetic insight with speech technology. In Barry, W., and Domelen, W. (Eds.), *Integrating Phonetic Knowledge with Speech Technology*. Kluwer, Dordrecht, The Netherlands.

Greenberg, S. (2006) A multi-tier framework for understanding spoken language. In Greenberg, S., and Ainsworth, W. A. (Eds.), *Listening to Speech: An Auditory Perspective.* Lawrence Erlbaum, Mahwah, NJ, pp. 411-434.

Greenberg, S. (2011). Speak, memory — Wherefore art thou, invariance? *The Journal of the Acoustical Society of America* 130, 2374. https://doi.org/10.1121/1.3654514.

Greenberg, S. (2018) Deep language learning. *Acoustics Today* 14(1), 19-27.

Greenberg, S., and Arai, T. (2004). What are the essential cues for understanding spoken language? *IEICE Transactions on Information Systems* 87-D, 1059-1070.

Greenberg, S., Carvey, H., Hitchcock, L., and Chang, S. (2003). Temporal properties of spontaneous speech–a syllable-centric perspective. *Journal of Phonetics* 31, 465-485. https://doi.org/10.1016/j.wocn.2003.09.005.

Hawkins, S. (2014). Situational influences on rhythmicity in speech, music, and their interaction. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 369(1658), 20130398. https://doi.org/10.1098/rstb.2013.0398.

Houtgast, T., and Steeneken, H. (1973). The modulation transfer function in room acoustics as a predictor of speech intelligibility. *Acta Acustica* 28, 66-73.

Kawahara, H. (2015). Temporally variable multi-attribute morphing of arbitrarily many voices for exploratory research of speech prosody. In Hirose, K., and Tao, J. (Eds.), *Speech Prosody in Speech Synthesis: Modeling and Generation of Prosody for High Quality and Flexible Speech Synthesis.* Springer-Verlag, Berlin. Heidelberg, Germany. https://doi.org/10.1007/978-3-662-45258-5_8.

Kotz, S. A., Ravignani, A., and Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences* 22, 896–910. https://doi.org/10.1016/j.tics.2018.08.002.

Leong, V., and Goswami, U. (2014). Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. *Frontiers in Human Neuroscience* 8, 96. https://doi.org/10.3389/fnhum.2014.00096.

Lubinas, C., Keitel, Oblesser, J., Poeppel, D., and Rimmele, J. (2022). Explaining flexible continuous speech comprehension from individual motor rhythms. *bioRxiv.* Available at https://www.biorxiv.org/content/10.1101/2022.04.01.486685v1.

Obermeier, C., Menninghaus, W., von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., and Kotz, S. A. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Frontiers in Psychology* 4, 10. https://doi.org/10.3389/fpsyg.2013.00010.

Oganian, Y., and Chang, E. F. (2019). A speech envelope landmark for syllable encoding in human superior temporal gyrus. *Science Advances* 5(11), 6279. https://doi.org/10.1126/sciadv.aay6279.

Park, H., Ince, R.A., Schyns, P. G., Thut, G., and Gross, J., (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Current Biology* 25, 1649-1653. https://doi.org/10.1016/j.cub.2015.04.049.

Pefkou, M., Arnal, L. H., Fontolan, L., and Giraud, A. L. (2017). θ-Band and β-band neural activity reflects independent syllable tracking and comprehension of time-compressed speech. *The Journal of Neuroscience* 37, 7930-7938. https://doi.org/10.1523/JNEUROSCI.2882-16.2017.

Pickett, J. M., and Pollack, I. (1963). Intelligibility of excerpts from fluent speech: Effects of rate of utterance and duration of excerpt. *Language and Speech* 6, 151-165.

Plomp, R. (2002). *The Intelligent Ear: On the Nature of Sound Perception.* Lawrence Erlbaum, Mahwah, NJ.

Poeppel, D., and Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience* 21, 322-334. https://doi.org/10.1038/s41583-020-0304-4.

Pollack, I., and Pickett, J. M. (1964). Intelligibility of excerpts from fluent speech: auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior* 3(1), 79-84. https://doi.org/10.1016/S0022-5371(64)80062-1.

Polyanskaya, L., and Ordin, M. (2015). Acquisition of speech rhythm in first language. *The Journal of the Acoustical Society of America* 138(3), EL199-EL204. https://doi.org/10.1121/1.4929616.

Ravignani, A., Dalla Bella, S., Falk, S., Kello, C. T., Noriega, F., and Kotz, S. A. (2019). Rhythm in speech and animal vocalizations: A cross-species perspective. *Annals of the New York Academy of Sciences* 1453, 79-98. https://doi.org/10.1111/nyas.14166.

Rosenberg, A. (2010). AutoToBI: A tool for automatic ToBI annotation. *Proceedings of Interspeech, 2010, 11th Annual Conference of the International Speech Communication Association*, Makuhari, Chiba, Japan, September 26-30, 2010, pp. 146-149.

Scherer, K. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40, 227-256.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., and Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Current Opinion in Neurobiology* 20, 172-176. https://doi.org/10.1016/j.conb.2010.02.010.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303-304. https://doi.org/10.1126/science.270.5234.303.

Silipo, R., and Greenberg, S. (1999). Automatic transcription of prosodic stress for spontaneous English. *Proceedings of the 14th International Congress of Phonetic Sciences,* San Francisco, CA, August 1-7, 1999, pp. 2351-2354.

Silverman, K. E., Beckman, M. E., Pitrelli, J. F., Ostendorf, M., Wightman, C. W., Price, P., Pierrehumbert, J. B., and Hirschberg. J. (1992). ToBI: A standard for labeling English prosody. *International Conference on Spoken Language Processing. ICSLP 1992,* Banff, AB, Canada, October 13-16, 1992.

Smith, Z. M., Delgutte, B., and Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature* 416, 87-90. https://doi.org/10.1038/416087a.

Stevens, K. (1998). *Acoustic Phonetics*. MIT Press, Cambridge, MA.

Tucker, B. V., and Wright, R. A. (2020). Speech acoustics of the world's languages. *Acoustics Today* 16(2), 56-64. https://doi.org/10.1121/AT.2020.16.2.56.

van Bree, S., Sohoglu, E., Davis, M. H, and Zoefel, B. (2021). Sustained neural rhythms reveal endogenous oscillations supporting speech perception. *PLoS Biology* 19(2), e3001142. https://doi.org/10.1371/journal.pbio.3001142.

van Wassenhove, V., Grant, K. W., and Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* 45, 598-607. https://doi.org/10.1016/j.neuropsychologia.2006.01.001.

Winn, M. B. (2018). Speech: It's not as acoustic as you think. *Acoustics Today* 14(2), 43-49.

Zhang, W., and Ding, N. (2017). Time-domain analysis of neural tracking of hierarchical linguistic structures. *NeuroImage* 146, 333-340. https://doi.org/10.1016/j.neuroimage.2016.11.016.

About the Author

Steven Greenberg
steven@siliconspeech.com

Silicon Speech
El Paso, Texas 79912, USA

Steven Greenberg received his AB in linguistics from the University of Pennsylvania (Philadelphia) and his PhD in linguistics (with a strong minor in neuroscience) from the University of California, Los Angeles (Los Angeles). His thesis was *Neural Temporal Coding of Pitch and Vowel Quality*. He served as a research professor in the Department of Neurophysiology, University of Wisconsin-Madison (Madison) and held faculty appointments in Linguistics and the International Computer Science Institute at the University of California, Berkeley (Berkeley), and the Technical University of Denmark (Kongens Lyngby). He has edited several books on the auditory bases of spoken language and has organized conferences and symposia on speech and hearing science and technology.

Pitch Perception in a Developing Auditory Brain

Bonnie K. Lau

Introduction

A harpist plucks the string of her harp and uses the enchanting sound of a musical pitch to captivate her audience. But how does the auditory system turn the vibration of a harp's string into a pitch that we perceive as music? And more importantly, why is this pitch so important for hearing?

Indeed, whether it is to understand the words spoken around us, for the enjoyment of music, or perhaps to hear the voice of a server in a noisy restaurant, we rely on pitch to navigate the complex acoustic environments around us every day. In music, melodies are composed of sequences of pitch changes. In speech, pitch in vowels contributes to their identification, whereas pitch in voices conveys information regarding emotion, attitude, and talker identity. In tonal languages such as Cantonese, pitch variations in words change their meaning. Pitch is also used to help segregate simultaneous sounds in noisy environments.

How pitch is encoded by the ear and the brain has been a topic of scientific debate for many decades (Yost, 2015). However, although pitch has been extensively studied in mature auditory systems, less is known about the development of pitch perception in humans. Indeed, if pitch is critical to navigating our noisy world, can infants perceive pitch once hearing begins? What do babies hear when the string of a harp is plucked?

Neurophysiological studies conducted in human and nonhuman primates show the involvement of the auditory cortex, one part of the auditory brain, in pitch processing (Bendor and Wang, 2005), Thus, infant pitch perception is particularly interesting because the auditory brain undergoes a protracted and extended period of development. Although infants show responses to sound in the third trimester of gestation (Birnholz and Benacerraf, 1983), significant immaturity in the auditory cortex is observed at birth and throughout the first year of life. As a consequence of this immaturity, it is hypothesized that early responses to sound are supported primarily by subcortical processing, with a transition to more adult-like cortical mechanisms after the first six months of life (Eggermont and Moore, 2012).

This article considers the studies of infant pitch perception in the context of what they reveal about auditory brain development and how sound is perceived with an immature brain. Previous articles in *Acoustics Today* have covered infant speech development (Vick, 2018) and other aspects of auditory brain development (Kanold, 2022), whereas this article focuses on the perception of pitch in sounds like speech and music. I first begin with a definition of pitch and the two primary ways that pitch is encoded by the ear, which form the basis of the place and temporal models of pitch perception.

What Is Pitch?

One commonly adopted definition of pitch is that it is an attribute of sound that can be ordered on a scale from low to high (American National Standards Institute, 2013). The pitch of a pure tone corresponds to the frequency of its single component in the spectral domain (**Figure 1B**) and to the period of the waveform in the time domain (**Figure 1A**). The musical pitch produced by a harp's string is an example of a complex pitch (**Figure 1C**). Musical notes or vowels in speech are examples of a harmonic complex, which consists of multiple frequency components that are all integer multiples of the fundamental frequency (f_0). Although a complex sound has many separate frequency components, the pitch of the sound is a unitary percept that corresponds to its f_0 (**Figure 1D**). However, it should
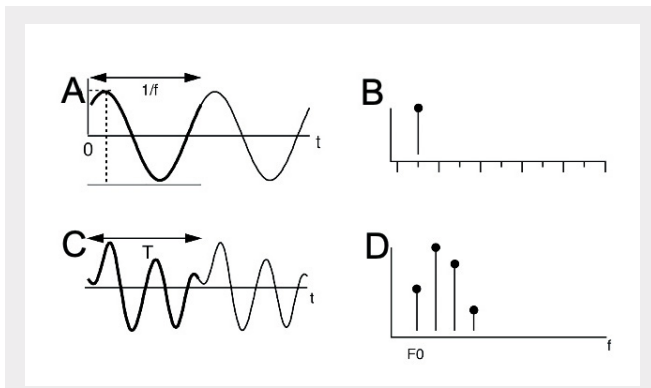
Figure 1. *A: time waveform of a pure tone where the period corresponds to its pitch (1/f). **B:** spectrum of a pure tone where the single frequency component corresponds to its pitch. **C:** time waveform of a complex tone where the period (*T*) corresponds to its pitch. **D:** spectrum of a complex tone where the fundamental frequency (*f₀*) corresponds to its pitch.*

be noted that pitch is defined perceptually as opposed to by a physical property of sound.

Place Versus Temporal Coding

There are two primary ways that the pitch is encoded in the ear that correspond with the spectrum and waveform of a sound. A rate-place code, which forms the basis of place models of pitch perception, represents the spectrum. When a sound enters the inner ear, it travels along the basilar membrane. Different frequencies in the sound maximally stimulate different regions along the basilar membrane, in effect doing a spectral analysis. Thus, the base of the basilar membrane is displaced by high-frequency components while the apex is displaced by low-frequency components. The frequency of a sound is encoded by the place of excitation on the basilar membrane, and the intensity of a sound is encoded by the firing rate of auditory nerve fibers at each place.

A temporal code, which forms the basis of temporal models of pitch perception, represents the time waveform. The temporal code relies on the tendency for auditory nerve fibers to fire at the same time during each cycle of vibration of the time waveform (phase locking). For most naturally occurring sounds, we would have access to both peripheral codes, but how these two codes are used by the ear and the brain to encode pitch is still a topic of active debate.

Early Infant Pitch Sensitivity

Despite significant immaturities in the auditory cortex, infants appear to respond to pitch in both speech and music. From the time they are born, infants show preference for their mother's voice over the voice of an unfamiliar female (DeCasper and Spence, 1986). Infants can discriminate a change in the frequency of pure tones (Wormith et al., 1975) and pitch contours in syllables (Karzon and Nicholas, 1989). They also show preference for high-pitched versus low-pitched singing (Trainor and Zacharias, 1998) as well as infant-directed speech, which is characterized by a high *f₀* and exaggerated pitch contours (Cooper and Aslin, 1990).

One of the challenging aspects of studying pitch perception is that oftentimes when the pitch in a sound changes, frequency and spectral changes also occur. However, many of these studies of early infant pitch perception did not differentiate between responses to frequency or spectral changes as opposed to pitch, and thus it is difficult to determine whether infants were indeed responding to pitch alone.

Observer-Based Psychophysical Procedure

Although it is difficult to determine what a preverbal infant hears, one approach that has a long history in both auditory and visual studies of infant perception is the observer-based psychophysical procedure (Werner, 1995). Sound discrimination can be measured with this method in infants from about 2 to 18 months of age. Sounds are presented to the infant through an insert ear tip or a speaker while the infant was sitting on a caregiver's lap inside a sound-attenuated booth (**Figure 2**). An assistant stays inside the booth to keep the infant attentive and facing the midline. No one hears what the infant is listening to; the caregiver is listening to music while the assistant is listening to the experimenter's instructions. The experimenter outside the booth starts each trial when the infant is facing the midline and attentive.

Infants are first conditioned to respond by pairing a large pitch change that they can perceive with reinforcers being shown, such as mechanical toys that move or a video that turns on. Common infant responses to these reinforcers include head turns or eye darts toward the mechanical animals, depending on the infant's age. The experimenter

Figure 2. *Observer-Based Psychophysical Procedure. A six-month-old infant is siting in a caregiver's lap listening to sounds via an insert ear tip (**left**). Infants are conditioned to respond by pairing a large pitch change with a video that turns on or mechanical animals that move (**right**).*

judges whether a change trial occurred based only on the infant's behavior.

Once an infant-experimenter pair has passed the conditioning phase of training, they move on to testing. During the experimental test phase, trials when a pitch change is played or no-change trials when sounds of the same pitch continue occur with equal probability. Different studies typically establish different criteria as evidence for discrimination. For example, infants may be required to have correct responses on four out of five consecutive change trials and four out of five consecutive no-change trials to demonstrate that they are responding to the sound changes.

Evidence of Central Pitch Extraction

Although infants demonstrate the ability to discriminate pure tones, given the immaturity of the auditory brain, it might be expected that they cannot perceive the pitch of a complex tone. Most models of pitch perception, regardless of whether it is a place or temporal representation, require a mechanism for pitch extraction. Thus, using the place code, the immature auditory system would need to resolve the harmonics of a complex tone, integrate the spectral information, and then extract the f_0. By contrast, using the temporal code, infant auditory nerve fibers would need to generate an accurate representation of the time waveform via phase locking and then extract the period.

There is evidence from past studies that pitch extraction requires processing in the brain to integrate information across the spectrum of a sound as opposed to encoding by the ear alone (Houtsma and Goldstein, 1972). Moreover, physiological studies suggest that an area outside the primary auditory cortex may be where pitch is extracted (Bendor and Wang, 2005). Given that until four months of age, only the most superficial layer of the auditory cortex contains mature axons (Moore and Guan, 2001), it would not be surprising if infants do not perceive complex pitch if a central mechanism for pitch extraction is required.

One method commonly used to control for responses to the f_0, or the pitch of a sound, is a classic phenomenon called *pitch of the missing fundamental*. The pitch produced by a harmonic complex tone is so strong that even if energy at the f_0 is missing, a pitch corresponding to the fundamental is still perceived. This phenomenon demonstrates that the auditory system must be able to extract information about the f_0 from the higher harmonics alone (see demonstration at tinyurl.com/ymrsfs2j). Requiring listeners to discriminate missing fundamental complexes also ensures that they are responding based on pitch as opposed to frequency or spectral changes because the f_0 is, in fact, not present in the tone.

Infants have demonstrated the ability to categorize missing fundamental complexes by pitch (Clarkson and Clifton, 1985). However, due to the nonlinear response of the cochlea, it is possible that discrimination of the missing fundamental pitch was based on combination tones produced by the cochlea (Pressnitzer and Patterson, 2001). One stimulus control that can limit the listeners' ability to rely on combination tones to perceive the pitch of the complexes is to use a noise band to mask the combination tones in the range of the missing fundamental. Indeed, infants as young as three months of age can categorize missing fundamental complexes by pitch, even in the presence of masking noise in the range of the missing fundamental (Lau and Werner, 2012). The results of these experiments suggest that central pitch extraction mechanisms are functional by this early age.

Unresolved Harmonics and the Temporal Code

By 6 months of age, infants can discriminate a change in frequency of a 4,000-Hz pure tone as well as adults do (Olsho et al., 1987). Discrimination ability at lower frequencies, however, continues to improve through

childhood and does not reach adultlike levels until adolescence (Maxon and Hochberg, 1982). One explanation for the difference between high and low frequencies is that the place code is used for high frequencies and develops rapidly, whereas the temporal code is used for low frequencies and develops more slowly.

In fact, many aspects of the development of temporal processing remain unknown (see Cabrera and Lau, 2022, for a review). Levi et al. (1995) recorded the frequency following response and the envelope following response to amplitude-modulated pure tones using electroencephalography (EEG), suggesting that temporal coding is functional in one month olds (Levi et al., 1995). Infants have also demonstrated the ability to perform discriminations that rely on temporal processing. For example, six month olds can discriminate speech contrasts when presented with processed syllables containing envelope cues but degraded temporal fine structure (Bertoncini et al., 2011).

However, when the temporal code is required due to limited spectral information for pitch, seven month olds demonstrate greater difficulty with pitch discrimination. The basilar membrane in the cochlea is often modeled as an array of band-pass filters known as auditory filters. Low-number harmonics fall in separate auditory filters and are separated by the cochlea and are thus called resolved harmonics. However, as the bandwidth of the auditory filters increases with increasing frequency, high-number harmonics fall within the same auditory filter and are referred to as unresolved harmonics because they cannot be separated by the cochlea. Unresolved harmonics thus rely on the temporal code for pitch because place-based information for pitch is not available.

Clarkson and Rogers (1995) found that more infants were able to discriminate complex tones composed of resolved harmonics compared with the number of infants that were able to discriminate unresolved harmonics. Butler et al. (2013) presented infants with high-pass filtered iterated ripple noise (IRN), a stimulus that also relies on the temporal code for pitch. To create IRN, a sample of noise is created, a delay is imposed on the noise, and then it is added back to the original noise iteratively. It also produces a weak pitch that increases in salience with the number of iterations. Butler et al. (2013) found that performance varied widely across infants and concluded that discriminating the pitch of IRN is difficult for infants. It

is possible that the infants' difficulty perceiving the pitch of unresolved harmonics and IRN in these two studies could be an indication of immature temporal pitch extraction. However, in a companion EEG study, Butler and Trainor (2013) did record a mismatch negativity to IRN pitch changes in infants despite the variability in behavioral performance they observed.

Furthermore, Lau and Werner (2014) found that both three- and seven-month-old infants were able to discriminate the pitch of unresolved harmonic complexes. This finding is consistent with past results showing that a temporal representation of pitch is available in the auditory nerve (Cariani and Delgutte, 1996). Moreover, this finding suggests that this temporal representation is functional in human infants by three months of age. Nevertheless, the variability in performance observed across studies suggests that temporal pitch is less salient for infants as it is for adult listeners.

High-Fidelity Pitch Discrimination

A hallmark of pitch perception is that adults can discriminate pitch with fine precision, an ability thought to be important for speech and music perception, as well as listening in complex acoustic environments. Many adult listeners can discriminate a less than 1% change in the f_0, and for those who cannot, rapid improvements in pitch discrimination can be observed even after brief periods of training (Micheyl et al., 2006).

Lau et al. (2021) compared discrimination of different degrees of pitch change in three and seven month olds as well as in musician and nonmusician adults to investigate the influence of cortical maturation on the acuity of pitch perception. They conducted a missing fundamental pitch categorization task to test the infants' and adults' ability to detect a change in the f_0 within a sequence of complex tones, each containing a random selection of consecutive harmonics, leading to random changes in timbre from tone to tone. Timbre is the perceptual attribute of sound that differentiates instruments (e.g., violin and guitar) or voices that are producing the same pitch and loudness. Surprisingly, both three and seven month olds performed as well as musician adults on this task, discriminating smaller changes in f_0 than the nonmusician adults.

This finding may be because the random variations in timbre presented in the tones used in this study

interfered with adult pitch discrimination but not that in infants. Stilp et al. (2010) have shown that adults are able to take advantage of statistical regularity in stimulus attributes to improve task performance and show rapid perceptual learning after a relatively brief exposure to covariations in sound features. As pitch and timbre often covary in natural sounds (Whalen and Levitt, 1995; Kitahara et al., 2005), this perceptual interference observed between pitch and timbre may be an efficient coding strategy. Thus, one explanation for the findings of Lau et al. (2021) is that infants have not learned the statistical covariation between pitch and timbre.

Nevertheless, the primary result of this study shows that accurate pitch and timbre discrimination can be achieved by infants as young as three months of age. Importantly, these findings imply that the fully mature auditory cortex is not required for accurate pitch discrimination, suggesting either that subcortical processing is sufficient for the f_0 and spectral coding or that the f_0 and spectral-peak discrimination is possible with an immature auditory cortex.

Melody Discrimination

One important aspect of pitch is that it that can be used to produce musical melodies. Although infants can discriminate complexes based on the f_0, they may not perceive melodic pitch in the same way as adult listeners. However, when presented with melodies composed of pure tones or two-component complexes, infants can detect changes to the melodies. Infants can detect a change in any position of a six-note melody (Trehub et al., 1985). The change in the melody can be in a different key or in the same key as the original melody (Trehub et al., 1984). Infants can even discriminate melodies composed of missing fundamental complexes (Lau et al., 2017). Interestingly, infants can detect changes in melodies belonging to scales from native Western scales as well as those from nonnative cultures (Lynch et al., 1990). Lynch et al. presented melodies based on native Western scales as well as nonnative Javanese pelog scales to American infants and adults and found that the infants were able to perceive the native and nonnative scales equally well, whereas the adults perceived the native scales better. These findings suggest that although infants are able to discriminate scales from different cultures, music perception is influenced by culturally specific listening experience by adulthood.

Musical Pitch Structure

Musical pitch structure is organized by several fundamental principles including consonance and dissonance, transpositional invariance, and tonal hierarchy. These principles appear to be perceived by infants from an early age and form the basis for other higher level pitch structures such as harmonic syntax, which does not appear until the childhood years (Trainor and Unrau, 2012).

Consonance and dissonance are considered fundamental organizing principles in musical pitch. Tones with f_0s in simple integer ratios such as the octave (2:1) or the perfect fifth (3:2) sound consonant, whereas tones with f_0s in complex integer ratios such as major seventh (15:8) tend to sound dissonant. Adult listeners from Western societies prefer consonance over dissonance, but there are divergent perspectives on the origin of this preference. One viewpoint is that consonance and dissonance arise from biological factors, whereas the opposing viewpoint is that it results from experiential factors (Weiss et al., 2020). For example, supporting the importance of experience, one study found that individuals from an Amazonian society that had minimal exposure to Western culture did not show a preference for consonance and dissonance (McDermott et al., 2016).

However, if infants perceive the distinction from a young age, that would be evidence to support that it is biological in origin. Indeed, Trainor and Heinmiller (1998) have shown that newborns and infants in the first few months of life look longer to consonant intervals than dissonant intervals, suggesting that they can discriminate between the two and that they prefer consonance over dissonance. Schellenberg and Trainor (1996) presented infants and adults with a sequence of consonant intervals and asked them to judge whether a test interval belonged to the sequence. Both infants and adults performed better on the task when the test interval was dissonant as opposed to consonant, suggesting that consonance has an influence on interval discrimination from a young age.

Consonance gives rise to the perception of tone chroma, also referred to as octave equivalence, the dimension of pitch that makes tones an octave apart sound similar. Demany and Armand (1984) found that by three months, infants show a perceptual equivalence for two pure tones forming an octave. They presented two melodies composed of pure tones to infants. The second melody

Figure 3. *Infant magnetoencephalography (MEG). Eight-month-old participant is being prepared for MEG recording including the placement of head position indicator (HPI) coils on a cloth cap to record head movement, electrocardiogram electrodes to record heartbeat, and digitization of HPI coils, cardinal landmarks, and 200 additional points on the head (**left**). A foam head bumper is placed on the infant to limit movement in the MEG dewar (**center**). Infant is awake and listening to sounds during MEG recording with a parent sitting in the chair next to them and an assistant in the booth showing them toys to keep them still and attentive (**right**). Photo courtesy of the University of Washington Institute for Learning & Brain Sciences (I-LABS), Seattle.*

consisted of tones shifted by an octave or another interval such as a seventh. Demany and Armand found that infants reacted less to transformed melodies consisting of pitch shifts by an octave than to shifts by larger or smaller intervals, suggesting that three month olds perceive tone chroma.

Finally, infants demonstrate the ability to perceive pitch contours in music. Relative pitch processing is critical to the perception of music because pitch relationships define melodies. Infants 5 to 10 months of age can recognize a familiar tune at any pitch after hearing it as little as 3 times. When tones are reordered or changed, infants perceive the melody as different, suggesting that infants can perceive relative pitch in melodies (Trehub et al., 1984).

Future Directions: What Is Happening in the Infant Brain?
The studies reviewed in this article suggest that for pitch, a fundamental aspect of sound, infants show accurate discrimination by about three months of age, despite significant immaturities in the auditory cortex at that age. Furthermore, infants' use of pitch for music perception seems to parallel that in adults in many ways.

The results of these studies also identify the many aspects of infant pitch perception that remain unclear. Future studies should further investigate infants' ability to use pitch while listening under noisy real-world conditions, and infants' ability to learn the statistical covariation between pitch and timbre as well as the development of pitch perception in infants with impaired pitch perception, such as those who use cochlear implants. Finally, neurophysiological studies of pitch perception can further our understanding of how pitch and sound in general is processed in the developing brain.

Magnetoencephalography (MEG) is one neuroimaging approach that allows for the recording of robust, temporally precise neural signals with high signal-to-noise ratios in infants as young as two months of age (**Figure 3**). Source localization techniques in MEG allow for the differentiation of subcortical and cortical sources and, perhaps most important for studying brain development, advanced techniques for tracking and correcting head movements that allow MEG to be recorded from infants who are awake and listening. **Figure 4** shows an example of infant neural responses to a pitch change, recorded with MEG; these signals can be studied in time and spectral domains and localized to their neural sources.
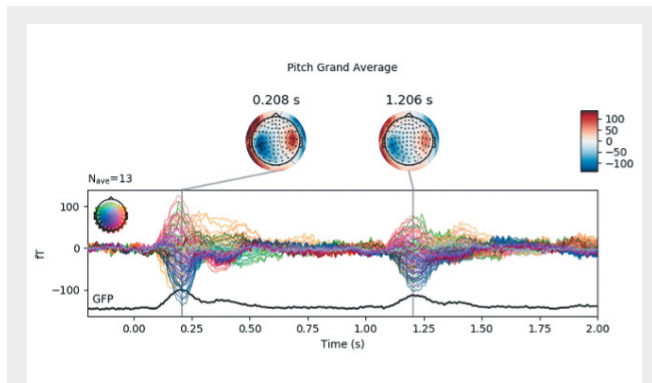
Figure 4. *Grand averaged (N$_{ave}$) event-related fields (magnetometers) from 13 six-month-old infants with topographic maps shown at the peak of onset (0.208 s) and acoustic change response (1.206 s) to a pitch change from 160 to 200 Hz.* **Bottom black line**, *global field power (GFP). Color bar units are femtotesla.*

With the rapid advances in neuroimaging technologies that can be used with infants, the hope is that we can further our understanding of this fundamental question in human auditory development: How, despite the protracted and extended developmental period of the auditory brain, infants demonstrate sophisticated sound-processing abilities from the time they are born.

References

American National Standards Institute (2013). *American National Standard Acoustical Terminology.* American National Standards Institute, New York, NY.

Bendor, D., and Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature* 436(7054), 1161-1165. https://doi.org/10.1038/nature03867.

Bertoncini, J., Nazzi, T., Cabrera, L., and Lorenzi, C. (2011). Six-month-old infants discriminate voicing on the basis of temporal envelope cues (L). *The Journal of the Acoustical Society of America* 129(5), 2761-2764. https://doi.org/10.1121/1.3571424.

Birnholz, J. C., and Benacerraf, B. R. (1983). The development of human fetal hearing. *Science* 222, 516-519.

Butler, B. E., and Trainor, L. J. (2013). Brief pitch-priming facilitates infants' discrimination of pitch-evoking noise: Evidence from event-related potentials. *Brain and Cognition* 83(3), 271-278. https://doi.org/10.1016/j.bandc.2013.09.002.

Butler, B. E., Folland, N. A., and Trainor, L. J. (2013). Development of pitch processing: Infants' discrimination of iterated rippled noise stimuli with unresolved spectral content. *Hearing Research* 304, 1-6. https://doi.org/10.1016/j.heares.2013.05.009.

Cabrera, L., and Lau, B. K. (2022). The development of auditory temporal processing during the first year of life. *Hearing, Balance and Communication* 1-11. https://doi.org/10.1080/21695717.2022.2029092.

Cariani, P. A., and Delgutte, B. (1996). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *Journal of Neurophysiology* 76(3), 1698-1716.

Clarkson, M. G., and Clifton, R. K. (1985). Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental. *The Journal of the Acoustical Society of America* 77(4), 1521-1528. https://doi.org/10.1121/1.391994.

Clarkson, M. G., and Rogers, E. C. (1995). Infants require low-frequency energy to hear the pitch of the missing fundamental. *The Journal of the Acoustical Society of America* 98(1), 148-154. https://doi.org/10.1121/1.413751.

Cooper, R. P., and Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development* 61(5), 1584-1595. https://doi.org/10.1111/j.1467-8624.1990.tb02885.x.

DeCasper, A. J., and Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development* 9(2), 133-150. https://doi.org/10.1016/0163-6383(86)90025-1.

Demany, L., and Armand, F. (1984). The perceptual reality of tone chroma in early infancy. *The Journal of the Acoustical Society of America* 76(1), 57-66. https://doi.org/10.1121/1.391006.

Eggermont, J. J., and Moore, J. K. (2012). Morphological and functional development of the auditory nervous system. In Werner, L., Fay, R. R., and Popper, A. N. (Eds.), *Human Auditory Development.* Springer, New York, NY, pp. 61-105.

Houtsma, A. J. M., and Goldstein, J. L. (1972). The central origin of the pitch of complex tones: Evidence from musical interval recognition. *The Journal of the Acoustical Society of America* 51(2B), 520-529. https://doi.org/10.1121/1.1912873.

Kanold, P. (2022). Listening to mom: How the early auditory experience sculpts the auditory cortex of the brain. *Acoustics Today* 18(1), 32-40. https://doi.org/10.1121/AT.2022.18.1.32.

Karzon, R. G., and Nicholas, J. G. (1989). Syllabic pitch perception in 2- to 3-month-old infants. *Perception and Psychophysics* 45(1), 10-14. https://doi.org/10.3758/BF03208026.

Kitahara, T., Goto, M., and Okuno, H. G. (2005). Pitch-dependent identification of musical instrument sounds. *Applied Intelligence* 23, 267-275. https://doi.org/10.1007/s10489-005-4612-1.

Lau, B. K., and Werner, L. A. (2012). Perception of missing fundamental pitch by 3- and 4-month-old human infants. *The Journal of the Acoustical Society of America* 132, 3874-3882. https://doi.org/10.1121/1.4763991.

Lau, B. K., and Werner, L. A. (2014). Perception of the pitch of unresolved harmonics by 3- and 7-month-old human infants. *The Journal of the Acoustical Society of America* 136, 760-767. https://doi.org/10.1121/1.4887464.

Lau, B. K., Lalonde, K., Oster, M.-M., and Werner, L. A. (2017). Infant pitch perception: Missing fundamental melody discrimination. *The Journal of the Acoustical Society of America* 141, 65-72. https://doi.org/10.1121/1.4973412.

Lau, B. K., Oxenham, A. J., and Werner, L. A. (2021). Infant pitch and timbre discrimination in the presence of variation in the other dimension. *Journal of the Association for Research in Otolaryngology* 22, 693-702. https://doi.org/10.1007/s10162-021-00807-1.

Levi, E. C., Folsom, R. C., and Dobie, R. A. (1995). Coherence analysis of envelope-following responses (EFRs) and frequency-following responses (FFRs) in infants and adults. *Hearing Research* 89(1-2), 21-27.

Lynch, M. P., Eilers, R. E., Oller, D. K., and Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science* 1(4), 272-276.

Maxon, A. B., and Hochberg, I. (1982). Development of psychoacoustic behavior: Sensitivity and discrimination. *Ear and Hearing* 3, 301-308.

McDermott, J. H., Schultz, A. F., Undurraga, E. A., and Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature* 535, 547-550. https://doi.org/10.1038/nature18635.

Micheyl, C., Delhommeau, K., Perrot, X., and Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research* 219, 36-47. https://doi.org/10.1016/j.heares.2006.05.004.

Moore, J. K., and Guan, Y. L. (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research in Otolaryngology* 2(4), 297-311.

Olsho, L. W., Koch, E. G., and Halpin, C. F. (1987). Level and age effects in infant frequency discrimination. *The Journal of the Acoustical Society of America* 82, 454-464. https://doi.org/10.1121/1.395446.

Pressnitzer, D., and Patterson, R. D. (2001). Distortion products and the perceived pitch of harmonic complex tones. *Physiological and Psychophysical Bases of Auditory Function* 97-104.

Schellenberg, E. G., and Trainor, L. J. (1996). Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: Tests of adult and infant listeners. *The Journal of the Acoustical Society of America* 100, 3321-3328. https://doi.org/10.1121/1.417355.

Stilp, C. E., Rogers, T. T., and Kluender, K. R. (2010). Rapid efficient coding of correlated complex acoustic properties. *Proceedings of the National Academy of Sciences of the United States of America* 107, 21914-21919. https://doi.org/10.1073/pnas.1009020107.

Trainor, L. J., and Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development* 21, 77-88. https://doi.org/10.1016/S0163-6383(98)90055-8.

Trainor, L. J., and Unrau, A. (2012). Development of pitch and music perception. In Werner, L., Fay, R. R., and Popper, A. N. (Eds.), *Human Auditory Development.* Springer, New York, NY, pp. 223-254. https://doi.org/10.1007/978-1-4614-1421-6_8.

Trainor, L. J., and Zacharias, C. A. (1998). Infants prefer higher-pitched singing. *Infant Behavior and Development* 21, 799-805. https://doi.org/10.1016/S0163-6383(98)90047-9.

Trehub, S. E., Bull, D., and Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development* 55, 821-830. https://doi.org/10.2307/1130133.

Trehub, S. E., Thorpe, L. A., & Morrongiello, B. A. (1985). Infants' perception of melodies: Changes in a single tone. *Infant Behavior and Development* 8(2), 213-223.

Vick, J. (2018). Advancing toward what will be: Speech development in infancy and early childhood. *Acoustics Today* 14(3), 39-47.

Weiss, M. W., Cirelli, L. K., McDermott, J. H., and Trehub, S. E. (2020). Development of consonance preferences in Western listeners. *Journal of Experimental Psychology: General* 149, 634–649. https://doi.org/10.1037/xge0000680.

Werner, L. A. (1995). Observer-based approaches to human infant psychoacoustics. In Klump, G. M., Dooling, B. J., Fay, R. R., and Stebbins, W. C. (Eds.), *Methods in Comparative Psychoacoustics.* Birkhäuser, Basel, Switzerland, pp. 135-146.

Whalen, D. H., and Levitt, A. G. (1995). The universality of intrinsic F_0 of vowels. *Journal of Phonetics* 23, 349-366. https://doi.org/10.1016/S0095-4470(95)80165-0.

Wormith, S. J., Pankhurst, D., and Moffitt, A. R. (1975). Frequency discrimination by young infants. *Child Development* 46, 272-275. https://doi.org/10.2307/1128864.

Yost, W. A. (2015). Psychoacoustics: A brief historical overview. *Acoustics Today* 11(3), 46-53.

About the Author

Bonnie K. Lau blau@uw.edu

Department of Otolaryngology—Head and Neck Surgery
University of Washington
CHDD Clinic Building, Box 357923
Seattle, Washington 98195, USA

Bonnie K. Lau is a research assistant professor at the University of Washington (Seattle). Her research investigates auditory brain and perceptual development and how that relates to language acquisition. She combines neurophysiological measures including electroencephalography and magnetoencephalography, psychophysics, and clinical assessment methods in her research.

How Our Brains Make Sense of Noisy Speech

Jonathan E. Peelle and Arthur Wingfield

Introduction

In August 1909, the French otolaryngologist Étiene Lombard came before a meeting of the French Academy of Sciences to report on the phenomenon that now bears his name. In the presence of a noisy background, speakers automatically talk louder in an attempt to maintain an effective signal-to-noise ratio (that is, keep their speech louder than the background noise so it can be heard) (Lombard, 1911). As anyone who has had a meal in a small noisy restaurant or a loud social gathering knows, the "Lombard effect" can quickly escalate, with all of the others in the room similarly attempting to speak louder than the other speakers. A surreptitious glance at your cell phone-based sound level meter can reveal startling levels of background babble.

Fortunately, when dealing with noise, whether in a restaurant or another potentially loud environment, the mammalian auditory system has evolved ways to extract a signal of importance (a partner's speech!) from the noisy surround. Although many of the mechanisms lie within the ear itself (Litovsky, 2012), the brain has also evolved amazing ways to enhance speech comprehension in the presence of noise. The focus of this article is the effects of noise on speech comprehension and the neural systems engaged when a listener is faced with this challenge.

During spoken communication, listeners need to determine the words produced by a talker so that they can understand the intended meaning. For many people, understanding speech in relaxed settings feels relatively automatic and effortless. However, this feeling is at odds with the remarkably complex feat our auditory system performs, namely, mapping a rapid and acoustically complex stimulus onto a set of learned categories (words). The average university graduate has a speaking vocabulary of tens of thousands of words and an even larger comprehension vocabulary. The listener's task, then, is to match the incoming acoustic input with the relevant mental representations (the "mental lexicon") of the words they know. What may be even more impressive is that this process must occur as the information is arriving at average speech rates of 140 to 180 words per minute, passing the ear, literally, at the speed of sound. Thus, much of our analysis of the speech signal lags behind the arriving acoustic input and must be carried out on a fading trace of the input in our short-term memory.

Given the time constraints governing speech perception, listeners become experts at using knowledge about speech and language, including what words are likely to come next given the preceding context, to aid understanding. So, for example, if you hear the sentence "I like cream and sugar in my…" you might expect the next word to be "coffee" or perhaps "tea," and this expectation will aid your understanding (coffee and tea will be recognized more quickly and accurately than "toffee"). Or, when listening to an unfamiliar talker, listeners typically adjust to this talker over time and become more efficient at understanding their speech. Even though the incorporation of acoustic and linguistic expectations usually happens without a listener's conscious awareness, on some level their brain is rapidly processing these types of information.

Although listening in quiet may feel relatively easy, listening in background noise can be noticeably challenging. Even when background noise does not completely drown out a talker, it can obscure sounds and make words ambiguous or unintelligible. And, if the background noise consists of other speech (as frequently happens in a coffee shop or restaurant), the content of the background speech can also be distracting (especially if it's interesting!). It is no wonder that listening to speech in

noise is a chief complaint among people seeking hearing health care.

Despite the challenges that background noise presents for speech perception, in many cases listeners are nevertheless able to correctly understand what a talker has said. Here, we explore the ways that listeners' neural systems within the brain deal with speech that is acoustically challenging. We use the term "acoustically challenging" speech to cover a broad range of challenges such as speech in background noise, speech heard in the midst of other talkers, understanding speech by listeners with hearing loss, and understanding the spectrally degraded sound delivered by a cochlear implant. We focus on studies suggesting that our brains need to "work harder" when listening to acoustically challenging speech than they do when listening to acoustically clear speech and the implications these findings have for everyday communication.

The Brain Systems Involved in Understanding Speech

Before exploring how listeners' brains respond to *challenging* speech, it will be helpful to review the core brain regions involved in understanding both sounds and speech. This pathway from the ears to the cortex is shown in **Figure 1**. Auditory information is passed from the cochlea (the inner ear) to the auditory nerve and then along several auditory nuclei (a nucleus is a collection of cells with similar function). These nuclei function, in part, to compare signals from the left and right ears to extract cues to spatial location, which can also aid in disentangling a target sound from background noise. Auditory information reaches the cortex in Heschl's gyrus (primary auditory cortex) on the top portion of the temporal lobe. From here, different brain regions are engaged depending on what is being heard (e.g., simple tones, speech, or environmental sounds) and the task being done.

In a side view, the four lobes of the brain (frontal, temporal, occipital, and parietal) are shown in the left hemisphere (**Figure 2a**). Investigators first learned about the brain regions involved in speech by observing patients who have had brain damage (e.g., due to a stroke) and who have, as a result, developed language difficulty (known as *aphasia*). The two most widely known types of aphasia are Broca's aphasia (caused by damage to the left frontal lobe and associated primarily with difficulty producing speech) and

Wernicke's aphasia (caused by damage to the left temporal lobe and associated primarily with difficulty comprehending speech). These conditions early on pointed toward an important role for the left hemisphere in understanding speech as well as highlighting contributions from both the temporal and frontal regions.

However, a great deal has been also learned from functional brain-imaging studies in which we are able to measure regional brain activity while people listen to speech. Among imaging approaches, functional magnetic resonance imaging (fMRI) has long been the most popular due to its wide availability (nearly every hospital or medical center has an MRI scanner) and spatial precision (Evans and McGettigan, 2017). fMRI takes

Figure 1. *Auditory processing pathways.* **Left:** *each region shown is a cross section of the brain at a different level of the auditory system.* **Right:** *side view of the brain. Sound enters the auditory system in the cochlea (inner ear) before proceeding up a complicated set of subcortical nuclei leading to the primary auditory cortex. Available at* osf.io/u2gxc, *under CC BY 4.0 Attribution 4.0 International license. See also Peelle and Wingfield (2016).*
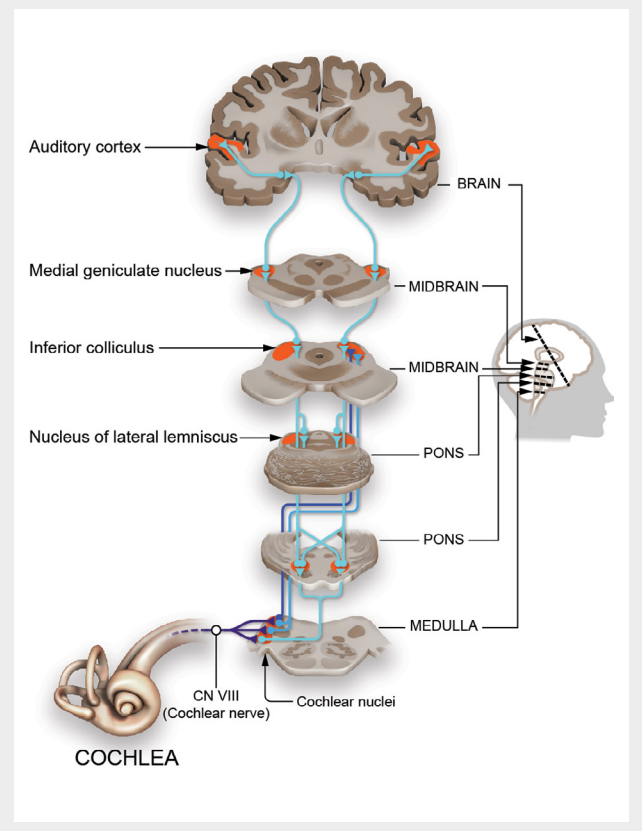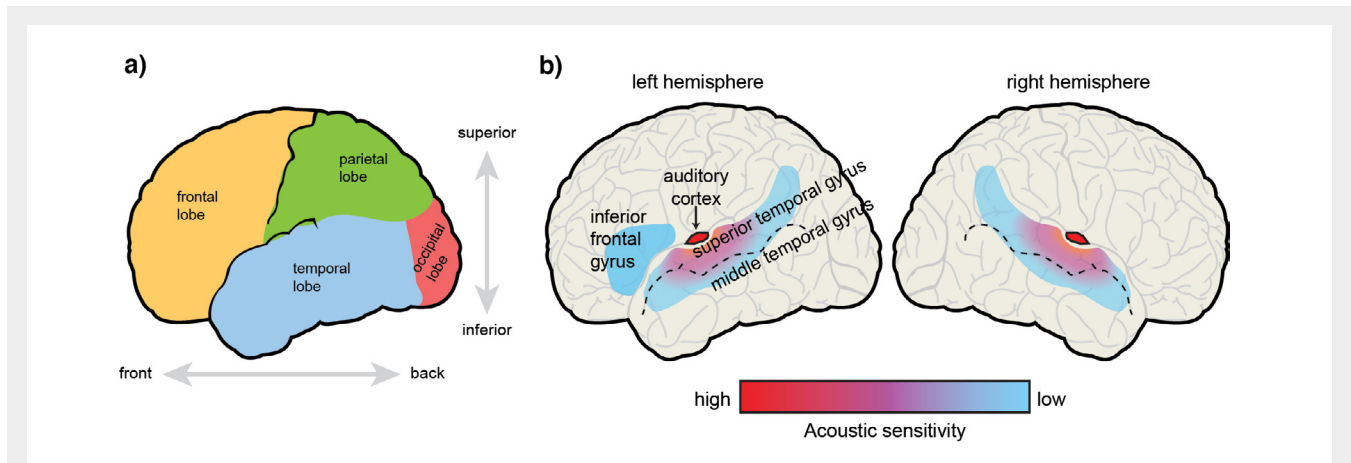
Figure 2. a: *Side view (also known as a lateral view) of the left hemisphere, with the four lobes of the brain indicated. Superior is used for structures closer to the top of a lobe or structure and inferior for those closer to the bottom.* **b:** *Views of the left and right hemispheres showing the cortical speech regions. From the auditory cortex, parallel pathways process speech with an increasing abstraction, reflecting increasingly less acoustic detail. Adapted from Peelle et al. (2010).*

advantage of the fact that areas of the brain increasing their relative level of activity draw an increase in the relative blood flow to that area to bring the oxygen needed to sustain this activity. The blood carrying oxygen (oxygenated hemoglobin) has different magnetic properties than deoxygenated hemoglobin, which can be detected by an MRI scanner.

Thus, although historically the language difficulties of people who lost functions due to brain damage gave us the broad outlines of regions in the left hemisphere important for language, fMRI and other modern brain-imagining techniques provided a more nuanced and complete view of core speech-processing regions (summarized in **Figure 2b**). What this modern work has shown is that when listening to single words, both the left and right temporal lobes are engaged. This includes not only the primary auditory cortex but also nearby regions on the superior temporal gyrus and middle temporal gyrus. (The surface of the human brain is not smooth but folded. A *gyrus* is a bump or a "mountain," and a *sulcus* is the crevice between bumps or a "valley.") Together, these regions of the left and right temporal lobes, encompassing the auditory cortex, superior temporal gyrus, and middle temporal gyrus, are responsible not only for processing the acoustic information in speech but also for linking the acoustic information to words and word meaning.

When listening to sentences and stories, the left inferior frontal gyrus also becomes active. Although the specific contribution of this frontal activity is debated (complicated by many smaller subdivisions of the inferior frontal cortex that seem to play distinct roles), many of these functions appear to relate to the rules for combining words to form a meaningful sentence. These grammatical rules are referred to as the syntax of a sentence. Regions of the left inferior frontal gyrus also respond to more complicated aspects of word meaning, such as understanding from the context whether "bark" might refer to the sound a dog makes or the covering on a tree. Thus, the core regions supporting speech understanding start with the auditory cortex and then continue to a more extended network concerned with various levels of language processing.

A key characteristic of human speech regions is that they are hierarchically organized; stages anatomically nearer the auditory cortex are more involved in processing the specific acoustic signatures of speech. For example, they respond differently depending on how speech is degraded (different kinds of background noise result in different patterns of brain activity). By contrast, activity in regions that are further away, such as in the frontal lobe, depends less on the acoustic details of speech and more on the informational content (e.g., whether speech is intelligible). These different components of the speech

network work in a coordinated way to translate the acoustic speech signal to its intended meaning.

How Young Adults with Normal Hearing Make Sense of Degraded Speech

Even young adults with good hearing must make sense of noisy speech. One way to study the cognitive consequences of noisy listening is through behavioral measures, such as asking people how well they remember what they have heard. Memory studies are useful for two reasons. First, in everyday life, we often would like to remember what we hear, and so studying the effect of acoustic challenge on memory has clear real-world implications.

Second, there is a clear theoretical framework that lets us use memory differences to understand cognitive processing. Specifically, such studies rely on the principle that the brain is limited in its computing capacity. Thus, if people are worse remembering noisy speech compared with easy-to-understand speech, it suggests that the presence of background noise increased cognitive demand during listening.

Of course, it is not very interesting to find that people have trouble "remembering" something if it was never understood in the first place. Thus, the clearest demonstrations of the effect of noise on memory occur when speech in noise is shown to be audible in an intelligibility check. In an early demonstration of this effect, Rabbitt (1968) presented listeners with lists of digits to recall. In one condition of his experiment, the second half of the list was always presented in clear, unprocessed, easy-to-understand speech. The first half of the list was sometimes presented in clear speech and at other times acoustically degraded speech. In this latter case, Rabbitt made sure that the words could still be understood (although with effort). Rabbitt found that when the first half of the list was degraded, listeners had trouble remembering the second half of the list. Acoustically, there is no reason for this change; the speech in the second half of the list was always clear and easy to understand. Rabbitt concluded that additional cognitive resources were required for the degraded speech to be understood, such that fewer resources were available for remembering subsequent information. Since this landmark demonstration, many other studies have shown that acoustic challenge interferes with memory, even when speech is intelligible (for a review, see Peelle, 2018).
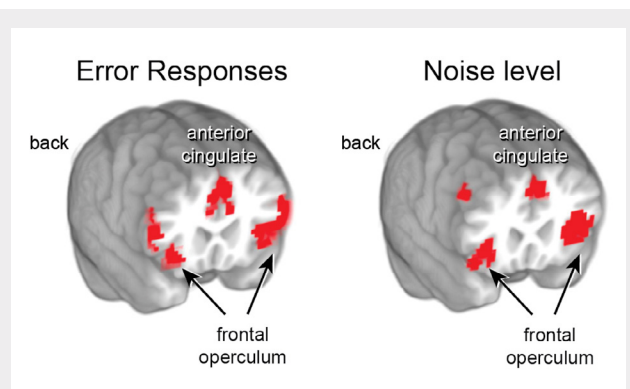
Figure 3. *Two images of a brain with a cross section cut to show deep structures (the very front of the brain has been removed). **Red**, regions of brain activity associated with error responses (**left**) and noise level (**right**). For young adults with normal hearing, the cingulo-opercular network composed of the anterior cingulate and bilateral frontal operculum is engaged during difficult listening, for both error responses and elevated noise conditions. These regions are not typically associated with speech processing in easy-listening situations. Adapted from Vaden et al. (2013).*

These behavioral studies linking acoustic challenge and memory suggest that a cognitive resource is used for both types of processing, and therefore increasing the cognitive demands of listening "steals" cognitive resources away from memory. Functional brain-imaging studies investigating how listeners process degraded speech are broadly consistent with this hypothesis, identifying regions of the frontal lobe not typically seen during "easy" speech perception that become more active when speech is acoustically challenging. In an elegant demonstration of this effect, Vaden and colleagues (2013) presented single words in background noise to listeners and had them repeat back each word as a measure of accuracy. The noise was difficult enough that some of the words were repeated correctly, whereas others were not. The results are shown in **Figure 3**, which includes two pictures of the brain highlighting different analyses. Following error trials, listeners showed increased activity in a network consisting of the anterior cingulate and frontal operculum, often called the cingulo-opercular network (**Figure 3**, *left*). **Figure 3**, *right*, shows many of the same regions but as a function of the noise levels of the speech. Activity in the cingulo-opercular network is associated with general task engagement and is frequently observed following errors on many kinds of tasks. A particularly

compelling aspect of Vaden and colleagues' study is that they found error-related activity in one trial predicted listening success in the *following* trial, consistent with activity in the cingulo-opercular network relating to task reengagement following a mistake.

The cingulo-opercular network is by no means the only brain response to challenging speech in normal-hearing young adults. However, the cingulo-opercular activity is notable for at least two reasons. First, anatomically, it unquestionably lies outside of the core speech network outlined above; the brain is clearly doing something different for degraded speech than it does for easy-to-understand speech. Second, the regions and response profile are consistent with domain-general processing that also goes along with behavioral observations. Understanding speech in noise requires cognitive resources not seen during easy-listening conditions.

Challenges to Speech Understanding in Adult Aging

Among older adults, hearing loss is one of the most common chronic medical conditions (Lethbridge-Cejku et al., 2004). Although age-related hearing loss is primarily a result of cochlear hair cell loss, especially those sensitive to high-frequency sounds, there can also be deterioration throughout the central auditory pathway, from the cochlear nucleus to the auditory cortex (Peelle and Wingfield, 2016). Adult aging is also accompanied by brain changes that affect the structure and network dynamics that carry cognitive function (Peelle and Wingfield, 2016; Anderson et al., 2018). Important consequences of these latter changes include a reduced capacity of working memory, a reduced ability to inhibit potential interference from concurrent stimuli, and a general slowing in a number of perceptual and mental operations. Despite these changes, barring neuropathology, speech comprehension generally remains well preserved in adult aging due in large part to older adults' effective use of linguistic and situational context.

There are, however, several circumstances that present a special challenge for the older listener. These include very rapid speech that places a demand on a slowed processing system, speech in which the meaning is expressed with complex syntax that places a heavy burden on working memory and, relevant to our present topic, speech heard in a noisy background. It is almost axiomatic that older adults have special difficulties in hearing speech in noise, often to a degree that would not be predicted from either auditory sensitivity (e.g., pure-tone thresholds) or the ability to hear speech in quiet (Anderson et al., 2018).

An underlying factor is older adults' reduced effectiveness in perceptually separating the target speech from background noise. This process is sometimes referred to as *auditory stream segregation* (Carlyon, 2004). Many consider these sound streams as "objects" that, once identified, can be selectively attended to (or ignored). Segregating auditory streams depends in large part on spatial cues but also on the physical features of the sounds. In everyday listening, background noise often fluctuates in intensity (amplitude "dips") or periods of brief silence ("gaps"), with there being a benefit to listeners when such dips or gaps are more frequent and of a longer duration. Older adults' speech recognition gains relatively less benefit from gaps and dips in the noise than those in young adults, although this is mitigated to some extent by the effective use of linguistic context.

As hinted, a special case arises when the "noise" consists of other speakers. The term "cocktail party problem" was coined by Cherry (1953) to refer to one's ability to attend to a single speaker while being unaware of the content of other talkers speaking simultaneously (see also Middlebrooks et al., 2017; Leibold et al., 2019). Following a single speaker in a cocktail party situation is more difficult for older adults than for young adults, and especially so for adults with even mild hearing loss. At least part of this decrement is due to interference at the cognitive level (e.g., due to distracting information). In one demonstration of this, we compared younger and older adults on their ability to repeat speech from a target speaker when overlaid by a second talker speaking meaningful English or a language unfamiliar to the listeners (Dutch). Consistent with long-standing findings, the young adults' performance was equally unaffected whether the concurrent speech was in English or Dutch. By contrast, however, when the competing speaker was speaking in meaningful English, the older adults had more difficulty, indicating that the content in the to-be-ignored speech could not be fully ignored (Tun et al., 2002). The fact that the interference was specific to the content of the noise is consistent with the importance of cognitive factors in the comprehension of speech in noise in older adults.
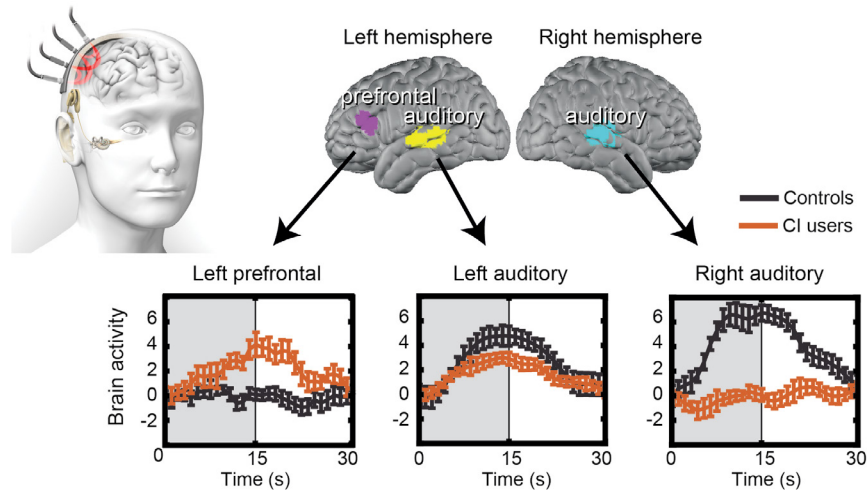
Figure 4. *Optical brain imaging provides a measure of regional brain activity like that obtained from functional magnetic resonance imaging (fMRI). Optical brain imaging was used to look at brain activity in listeners with cochlear implants (CIs) while they listened to spoken words. Three regions were looked at specifically: left auditory cortex, right auditory cortex, and left prefrontal cortex. Listeners with CIs showed different patterns of activity compared with listeners with good hearing, most notably increased activity in the prefrontal cortex. Adapted from Sherafati et al. (2022). Functional near-infrared spectroscopy (fNIRS) figure (**top left**) available at* osf.io/t8bxe, *under CC BY 4.0 Attribution 4.0 International license.*

Challenges to Speech Understanding from Cochlear Implants

In some cases of deafness, a cochlear implant can be used to restore hearing by electrically stimulating the auditory nerve (Goupell, 2015; Wilson, 2019). The clarity of speech processed through a cochlear implant, however, is sharply degraded compared with what the brain receives from normal (acoustic) hearing. As a result, listeners with cochlear implants frequently find speech understanding very effortful.

One way to measure the cognitive challenge experienced by listeners with cochlear implants is to measure brain activity during listening. However, the cochlear implant hardware presents specific challenges. In particular, patients with cochlear implants typically can't have an MRI, and the hardware also creates electrical and magnetic signals that interfere with other forms of brain imaging. One solution to these challenges has been to use optical brain imaging, commonly referred to as functional near-infrared spectroscopy (fNIRS). In fNIRS, experimenters shine a light on the skull. Some of the light gets absorbed and some travels through the head, into the brain, and back to the surface, where it can be measured. With knowledge about the light wavelengths that

are absorbed by oxygenated and deoxygenated hemoglobin, researchers can estimate regional blood flow in the brain that is strongly correlated with brain activity. Optical brain-imaging arrays vary in the number of measurements they provide and thus in how accurate spatial localization can be.

Recently, high-density diffuse optical tomography, a form of optical brain imaging, was used to measure brain activity in listeners with cochlear implants while they listened to spoken words (Sherafati et al., 2022). The pattern of activity produced is summarized in **Figure 4**. Compared with controls, adult listeners with cochlear implants showed greater activity in the dorsolateral prefrontal cortex (part of the frontal lobe). These findings are notable because this part of the brain does not seem to be regularly engaged in speech comprehension. Instead, the dorsolateral prefrontal cortex is usually associated with executive tasks such as attention, decision making, and some forms of short-term memory. The implication of these findings is that because of the unclear acoustic signal, the brains of listeners with cochlear implants need to work harder to make sense of what they are hearing. This additional cognitive effort may interfere with higher level understanding or make it harder to remember what

has been heard. However, it is worth noting that uncovering the brain systems supporting speech in listeners with cochlear implants is an active and relatively new area of research, and we expect our understanding to substantially evolve over the coming years.

Challenges to Speech Understanding from Face Masks

Although different types of face coverings have long been used in medical, industrial, and social contexts, widespread public health guidance regarding the benefits of face masks during the Covid-19 pandemic brought public awareness about face masks and associated communication challenges to a new level. Face masks challenge speech processing in at least two ways. First, the mask material partially blocks sound transmission, especially at higher frequencies, making speech not only potentially softer but obscuring specific speech cues. Second, opaque face masks prevent access to visual speech information from a talker's mouth, which is often relied on by listeners. The use of visual speech information is especially important for listeners with hearing loss or with a cochlear implant.

To look at how different kinds of face masks affected speech processing, people were asked to listen to sentences spoken by a talker with and without a mask (Brown et al., 2021). The sentences could be in quiet or in noise and were spoken with different face masks: a cloth mask without a filter, a cloth mask with a filter, a surgical mask, and a consumer transparent face mask, (**Figure 5a**). After each sentence, the people were asked to report the words they heard as a measure of their intelligibility and also to rate how difficult it was to understand the speech (as a measure of cognitive effort). Differences in performance were found depending on what kind of mask the speaker wore (**Figure 5b**). The surgical mask had the best performance, and the cloth mask with a filter and transparent mask performed the most poorly. Importantly, there were differences not only in speech intelligibility but also in the perceived effort associated with listening.

It is important to emphasize that Brown et al. (2021) tested a single type of clear face mask with listeners who reported normal hearing. It is very likely that for some listeners, visual speech information is crucial for effective communication; it is also likely that better clear masks exist rather than the one we tested. The data simply indicate that a clear
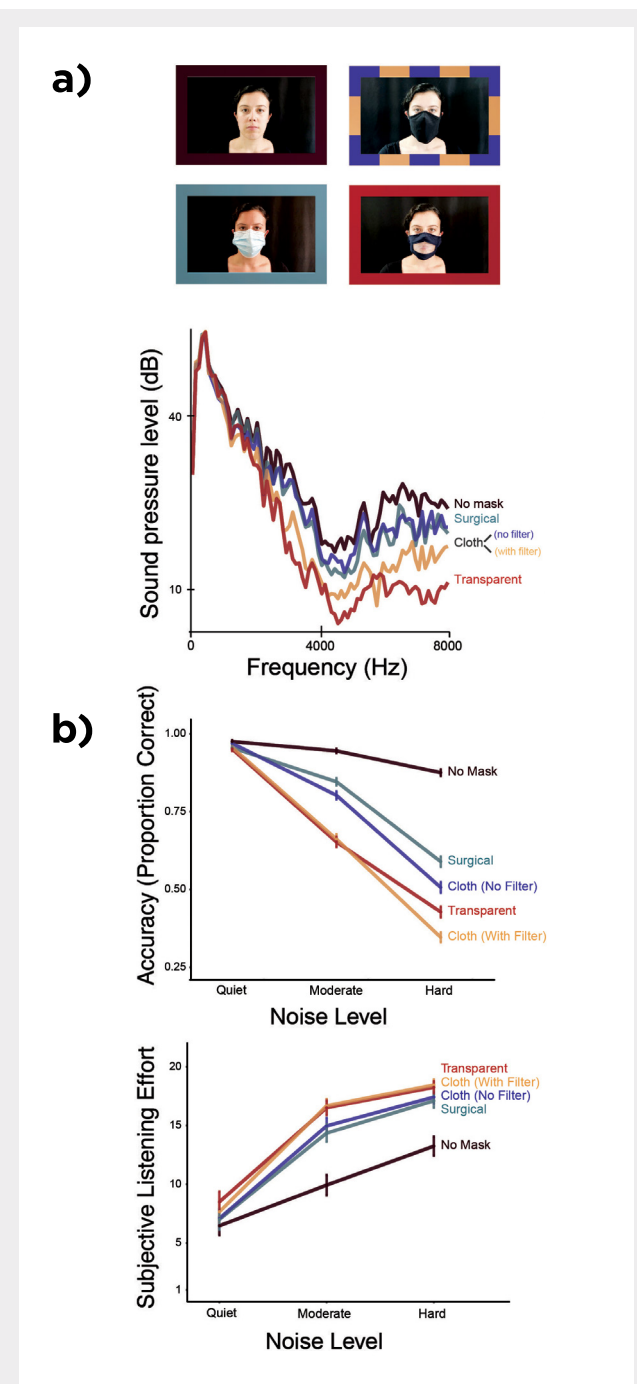
Figure 5. a: *Face masks tested (**top**) and the power at different frequency ranges (that is, the long-term average spectrum) of speech produced by these masks (**bottom**). The differences in the long-term average spectra indicate that different masks affect speech energy differently.* **b:** *Effect of face masks is larger in noise than in quiet, assessed both in terms of speech intelligibility (**top**) and how difficult listeners perceived the task to be, or subjective effort (here data from older adults; **bottom**). Adapted from Brown et al. (2021).*

mask is not *always* better for spoken communication and that other factors must be considered. And, indeed, in this study, N95/KN95 masks, which provide superior protection, were not evaluated.

How to Protect Hearing at Any Age and the Importance of Hearing Health Education

Hearing protection and monitoring ambient sound levels have seen increasing use in industrial settings in the United States and many other countries. An often-expressed concern, however, is potential effects on hearing related to the persistent high sound levels at many concert venues and with personal music players, especially among young adults.

Concern about sound exposure is heightened by studies showing that many young adults are unaware that they are beginning to have a hearing loss. The implications of these findings relate directly to our earlier mention that successful recognition of noise-masked speech comes at the cost of resources that would otherwise be available for encoding the words in memory. In an exploratory study, data obtained with university undergraduates who varied in hearing acuity were examined. All the undergraduates fell within a range typically considered in clinical audiology as normal hearing. The task involved the everyday experience of interpreting the meaning of spoken sentences. When the sentences expressed their meaning with a relatively simple syntax, interpretation accuracy was equivalently high for those at the higher and lower ranges of normal hearing. However, when sentences were presented that expressed their meaning with a more complex syntax, those with better hearing were more accurate than those with poorer hearing (Ayasse et al., 2019).

In addition to such behavioral effects, relatively small differences in hearing acuity among adults with normal or near-normal hearing manifest neural differences during sentence comprehension as well. Using fMRI, it was found that successful comprehension of sentences by individuals with poorer hearing was associated with greater recruitment (increased brain activity) in a right anterior middle frontal gyrus component of the frontoparietal attention network (Lee et al., 2018). These results demonstrate that even modest differences among individuals with clinically normal hearing affect the brain's response in the form of an increase in neural engagement of a non-sentence-specific component of the neural network to support successful sentence comprehension.

Thus, even slight differences in hearing acuity can have effects, however subtle, on everyday speech communication. At a practical level, these data argue for routine baseline hearing testing, even for young adults who currently have normal hearing. They also add to the growing support for an increased awareness of risks to hearing from extreme or prolonged noise exposure and, with it, increased use of hearing protection and noise reduction strategies.

Conclusions

Despite the challenges that a noisy acoustic signal presents for speech understanding, listeners' brains are able to engage additional cognitive systems to counter or at least mitigate the effects of noise on speech comprehension. However, doing so is not free but comes with a cognitive "cost": the increased processing needed for understanding speech may interfere with other mental activities, such as remembering what has been heard. Protecting hearing and, if needed, obtaining hearing health care, such as hearing aids, may thus have direct benefits for overall cognition.

At a broader level, studies such as those described in this brief review illustrate the general principle that the brain maintains stable functions (in this case, speech understanding) despite perturbations in the input (in this case, noise interference) by the flexible engagement of supporting neural networks. Speech comprehension, whether clear speech or speech in noise, must thus be understood as a dynamic and flexible interaction of the sensory, cognitive, and neural systems. The future of research and clinical practice in this area, we suggest, lies in our understanding of these interactions.

Acknowledgments

References

Anderson, S., Gordon-Salant, S., and Dubno, J. (2018). Hearing and aging effects on speech understanding: Challenges and solutions. *Acoustics Today* 14(4), 10-18.

Ayasse, N. D., Penn, L. R., and Wingfield, A. (2019). Variations within normal hearing acuity and speech comprehension: An exploratory study. *American Journal of Audiology* 28, 369-375. http://doi.org/10.1044/2019_AJA-18-0173.

Brown, V. A., Van Engen, K. J., and Peelle, J. E. (2021). Face mask type affects audiovisual speech intelligibility and subjective listening effort in young and older adults. *Cognitive Research: Principles and Implications* 6, 49. https://doi.org/10.1186/s41235-021-00314-0.

Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences* 8, 465-471.

Cherry, E. C. (1953). Some experiments on the recognition of speech with one and with two ears. *The Journal of the Acoustical Society of America* 25, 975-979.

Evans, S., and McGettigan, C. (2017). Comprehending auditory speech: previous and potential contributions of functional MRI. *Language, Cognition and Neuroscience* 32, 829-846. https://doi.org/10.1080/23273798.2016.1272703.

Goupell, M. J. (2015). Pushing the envelope of auditory research with cochlear implants. *Acoustics Today* 11(2), 26-33.

Lee, Y., Wingfield, A., Min, N. E., Kotloff, E., Grossman, M., and Peelle, J. E. (2018). Differences in hearing acuity among "normal-hearing" young adults modulate the neural basis for speech comprehension. *eNeuro* 5, ENEURO.0263-17.2018. http://doi.org/10.1523/ENEURO.0263-17.2018.

Leibold, L. J., Buss, E., and Calandruccio, L. (2019). Too young for the cocktail party? *Acoustics Today* 15(1), 37-43.

Lethbridge-Cejku, M., Schiller, J. S., and Bernadel, L. (2004). Summary health statistics for U.S. adults: National Health Interview Survey, 2002. *Vital Health Statistics* 10, 1-151.

Litovsky, R. (2012). Spatial release from masking. *Acoustics Today* 8(2), 18-25.

Lombard, E. (1911). Le signe de l'élévation de la voix. *Annales des Maladies de l'Oreille et du Larynx* XXXVII(2), 101-109.

Middlebrooks, J. C., Simon, J. Z, Popper, A. N., and Fay, R. R. (2017). *The Auditory System at the Cocktail Party*. Springer, Cham, Switzerland.

Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and Hearing* 39, 204-214.

Peelle, J. E., and Wingfield, A. (2016). The neural consequences of age-related hearing loss. *Trends in Neurosciences* 39, 486-497.

Peelle, J. E., Johnsrude, I. S., and Davis, M. H. (2010). Hierarchical processing for speech in human auditory cortex and beyond. *Frontiers in Human Neuroscience* 4, 51. http://doi.org/10.3389/fnhum.2010.00051.

Rabbitt, P. M. A. (1968). Channel capacity, intelligibility and immediate memory. *The Quarterly Journal of Experimental Psychology* 20, 241-248.

Sherafati, A., Dwyer, N., Bajracharya, A., Hassanpour, M. S., Eggebrecht, A. T., Firszt, J. B., Culver, J. P., and Peelle, J. E. (2022). Prefrontal cortex supports speech perception in listeners with cochlear implants. *eLife* 11, e75323. https://doi.org/10.7554/eLife.75323.

Tun, P. A., O'Kane, G., and Wingfield, A. (2002). Distraction by competing speech in young and older adult listeners. *Psychology and Aging* 17, 453-467.

Vaden, K. I., Jr., Kuchinsky, S. E., Cute, S. L., Ahlstrom, J. B., Dubno, J. R., and Eckert, M. A. (2013). The cingulo-opercular network provides word-recognition benefit. *Journal of Neuroscience* 33, 18979-18986.

Wilson, B. S. (2019). The remarkable cochlear implant and possibilities for the next large step. *Acoustics Today* 15(1), 53-61.

About the Authors

Jonathan E. Peelle
j.peelle@northeastern.edu

Center for Cognitive and Brain Health
Department of Communication
Sciences and Disorders and
Department of Psychology
Northeastern University
Boston, Massachusetts 02115, USA

Jonathan E. Peelle is an associate professor in the Center for Cognitive and Brain Health, Department of Communication Sciences and Disorders, and the Department of Psychology, Northeastern University (Boston, Massachusetts). He earned his PhD in neuroscience from Brandeis University (Waltham, Massachusetts; with Arthur Wingfield) and received subsequent training in human brain imaging at the University of Pennsylvania (Philadelphia) and the MRC Cognition and Brain Sciences Unit, Cambridge University (Cambridge, United Kingdom). Research in his lab, funded primarily by the National Institute for Deafness and Other Communication Disorders, is focused on understanding successful communication across the life span using behavior, physiology, and human brain imaging.

Arthur Wingfield
wingfiel@brandeis.edu

Department of Psychology and Volen
National Center for Complex Systems
Brandeis University
Waltham, Massachusetts 02454, USA

Arthur Wingfield is a professor of psychology and neuroscience at Brandeis University, Waltham, Massachusetts. He holds a master's degree in speech pathology and audiology from Northwestern University, Evanston, Illinois, and a doctorate in experimental psychology from the University of Oxford, Oxford, United Kingdom. His honors have included the Baltes Distinguished Research Achievement Award for his work on aging and speech perception and two MERIT Awards from the National Institute on Aging. He has served on review panels and task forces for the National Institutes of Health, the American Academy of Audiology, and the National Academy of Sciences. His research is supported by the National Institute on Deafness and Other Communication Disorders and the Stephen J. Cloobeck Research Fund.

Biotremology: Tapping into the World of Substrate-Borne Waves

Louise Roberts and Kyle Wickings

Introduction

The gargantuan sandworms in Frank Herbert's novel *Dune* and the worm-like creatures in the horror film *Tremors* may be fictional, but aspects of their sensory biology are close to real life (see the picture of *Dune* at bit.ly/3PQdixl; the *Tremors* movie clip at bit.ly/3t4Xn4t). These monsters of our imaginations utilize a sensory mode that humans rarely think about unless we live near a seismic fault line. In fact, when considering the environmental information around us, the sensory modes that spring to mind first are likely hearing, vision, and smell.

However, the humans in *Dune* were aware of another mode. Using "thumpers" to hit the ground rhythmically, the characters were able to attract sandworms. These creatures were sensitive to substrate-borne vibrations. In real life, we rarely notice that the substrates in our world are vibrating. Yet the use of vibrations in surfaces is hypothesized to be one of the oldest animal communication forms (Endler, 2014). Like the soundscape around us (Pijanowski et al., 2011), the vibrational landscape is termed the "vibroscape," a term that is applicable to both aquatic (Roberts and Elliott, 2017) and terrestrial (Šturm et al., 2021) systems. And it is only with the recent advance of sensor technologies and computer processing that scientists have started to familiarize themselves with this world (Hill and Wessel, 2016). This article explores biological vibrations selectively in two systems (our own backyard and the seashore) and discusses ways that humans are contributing to the vibrational landscape.

Biotremo- ...What?

Let us start at the beginning. The study of the biological use of vibrational waves, known as "biotremology," is a relatively new discipline that was first outlined in 2016 (Hill and Wessel, 2016). Several different wave types can travel through solids and substrates. Most biotremology studies involve surface waves, either Rayleigh or Bending waves, that travel in the boundaries between two media. The term "substrate-borne vibration" refers to these surface waves that travel in any substrate on which an animal resides such as hard or soft sediment, the water surface, or a plant. Sound waves may also travel in substrates, but biotremology differs from "bioacoustics" in that it refers to animals that use specialized vibrational receptors rather than pressure sensitive "ears" (Hill and Wessel, 2016).

In the animal kingdom, vibrations are used extensively for communication, including for parental care, foraging, detection of environmental cues, recognition, and predator-prey interactions (reviewed in Hill, 2008). Over 30 species of mammals actively produce vibrations, from kangaroo rats (e.g., *Dipodomys* spp.) drumming their hindfeet to elephant (e.g., *Loxodonta africana*) calls propagating into the ground. In the invertebrates, hundreds of thousands of species produce vibrations, including fruit flies (*Drosophila melanogaster*) and stinkbugs (e.g., *Nezara viridula*) tremulating (trembling; see video of fruit flies at youtu.be/519_XzM970s; stinkbugs at youtu.be/Q39C5I9L7mI), pill bugs (*Armadillo officinalis*) stridulating (rubbing body parts together), and crustaceans (Ocypodidae) drumming.

Animals detect incidental vibrations too, for locating prey and avoiding predators. For example, scorpions (*Paruroctonus mesaensis*) use vibrations to detect their arthropod prey (Brownell and Farley, 1979). Tree frog embryos (*Agalychnis callidryas*) hatch early after sensing vibrations from approaching predatory snakes (Warkentin, 2005). Ground-dwelling animals also detect abiotic vibratory cues such as thunder, which may be sensed in the substrate up to 1 km or more from the source (O'Connell-Rodwell et al., 2001), and even from tsunamis (1-100 Hz) as demonstrated in elephants (e.g., *Loxodonta africana*).

Indeed, vibration detection may be advantageous when other sensory modes (e.g., sound) are masked, such as by strong winds. However, because wind, temperature, and background noise affect acoustics transmission, substrate type and substrate composition also affect vibration transmission.

Given the extent of vibration use, why is this research area so far behind that of airborne communication? The explanation is twofold. First, humans do not sense vibrations as well as they are detected by other species. Our pressure-detecting ears bias us toward the soundscape. Thus, although airborne sounds have been meticulously documented, the realm of substrate-borne stimuli was overlooked for decades as a possibility for communication (Hill, 2008). Furthermore, substrate heterogeneity was thought to prevent this sort of communication.

Second, until sensor technologies and computation improved, the ability to detect vibrations was simply not available to scientists. Now, however, we can record vibrations by utilizing laser vibrometers, accelerometers, geophones, and piezoelectric sensors. Thus, we are aware of the vibrational world.

Vibrations Above and Below the Lawn

The realm of vibrations can be explored beneath the home lawn (**Figure 1**). Here, beneficial earthworms tunnel through the ground, improving the soil structure and recycling nutrients by feeding on organic matter.

Bait collectors use various methods to remove earthworms from the soil (see youtu.be/3ILoGcSxCAY) (Raboin, 2021). These methods include repetitious scraping of a metallic object against a stake (worm "grunting"), the twanging and moving of a garden fork ("charming"), repeated foot stamping, and electric motors applied to the ground (Catania, 2008). In fact, the annual Worm Gruntin' Festival in Sopchoppy, Florida, is in its 20th year!

What do these techniques have in common? Darwin (1882) stated that "earthworms are indifferent to shouts, nor did they [notice] the deepest and loudest tones of a bassoon," indicating a lack of sound reception. However, after placing pots of worms on the keys of a piano, Darwin observed a sensitivity to vibrations. In fact, earthworms have multiciliate sensory cells along the body surface and can detect tactile stimulation across
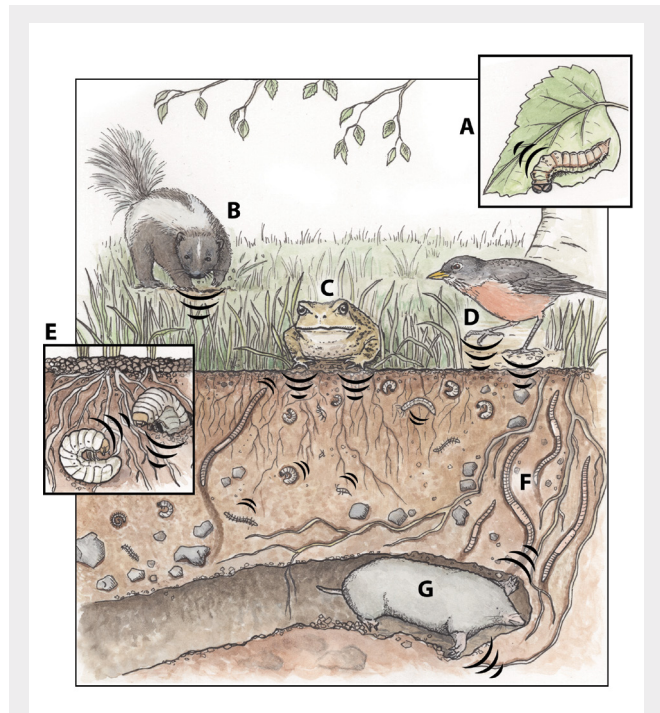
Figure 1. *Cross section of a North American lawn, with vibrational users and producers. Caterpillars produce leaf vibrations (**A**). Skunks are known for "stomping" behavior and for localizing grub prey with great accuracy (**B**). Toads "tremble" the toes (**C**) and birds "paddle" (**D**) during foraging bouts. Beetle larvae have been recorded stridulating (**E**). Earthworms (**F**) flee from the vibrations of their mole predators (**G**). Paintings by Ann Sanderson, © 2022, Ann Sanderson, all rights reserved; see www.annsciart.com.*

the whole body, making vibration reception highly likely (Laverack, 1960). Vibrational measurements of human-produced "grunts" indicate that they are broadband low frequency (<500 Hz) (Mitra et al., 2009). The bait collectors then are taking advantage of the earthworms' vibrosensitivity.

In tests involving mimics of rainfall (<500 Hz) and of the eastern American mole (*Scalopus aquaticus*) digging (<1,000 Hz), Catania (2008) found that earthworms (*Diplocardia mississippiensis*) were responding to the vibrations of their predator. Fossorial mammals produce these incidental vibrations (e.g., from digging) but also actively produce vibrational signals during territorial defense and competition (Mason and Narins, 2011). Worm grunters have unknowingly been mimicking mole vibrations.

Other animals have similar strategies to trigger prey emergence. The wood turtle (*Clemmy insculpta*) performs bouts of "stomping" to lure various segmented worms to the surface (see youtu.be/YAPTHDrAQw8). The turtle rocks from one foot to the other in a rhythmic trampling motion while periodically checking for emerged prey in the leaf litter (Kaufman, 1989). Birds also paddle the ground, discussed in **From the Lawn to the Seashore**. In anurans, a behavior called "toe twitching" has been described where the toes are vibrated, creating a "toe lure" of both visual and vibrational cues (Gridi-Papp and Narins, 2010). The brown marsh frog (*Rana baramica*), for example, vibrates the longest middle toe of each foot to attract prey (Grafe, 2008).

Above the lawn in the tree canopies, vibrational use is also widespread. Here a number of caterpillars, such as those of the moth family Drepanidae, actively produce leaf vibrations (Yack et al., 2001). The masked birch caterpillar (*Drepana arcuata*) builds silk nests on the leaf and defends the nest with signals consisting of "rasps" and "drums." Resident caterpillars are silent when solitary but use the open mandibles to strike the leaf and/or use the abdomen to scrape the surface to produce repetitive signals (see the supplementary videos at bit.ly/3t1GN5k).

Caterpillars do not have sound-sensing organs, so it is thought that the vibrational component is the primary informational source. The signals are elicited when the intruder is only 2-3 cm away, yet the vibrations travel much further and are stronger than required to signal to the intruder alone. This has led researchers to hypothesize that the signal is meant to attract the attention of predatory birds, putting the intruder at risk until the contest has ceased (Yack et al., 2001).

Leaf vibrations can also provide cues to other organisms. The incidental vibrations produced by caterpillars when chewing (for an example, see youtu.be/oEGIL9T73cQ) are detected by the plant itself, triggering a chemical defense response (Appel and Cocroft, 2014).

In the upper soil layers, beetle larvae can be found feeding within the turf thatch (an organic layer of roots, stems, and shoots), whereas other larvae are found deeper in the soil, feeding on decaying wood and tree roots. Stag beetle (*Lucanus cervus*), rose chafer (*Cetonia aurata*), and lesser stag (*Dorcus parallelipipedus*) larvae stridulate during active periods, such as when feeding, producing signals

Figure 2. *An adult slug beetle (*Lucanus cervus*). Larvae of the stag beetle are known to stridulate. Photo by Lucos Houska.*

in the region of 0.4-3 kHz (**Figure 2**) (Harvey et al., 2011). Similarly, the common (*Melolontha meloionth*) and forest (*M. hippocastani*) cockchafers also stridulate, with peak frequencies of 1.7 to 3 kHz that vary in duration depending on species (Görres and Chesmore, 2019).

The ecological meaning of these larval signals is unknown, but the use of vibration for communication venues such as reproduction, competition, and predator avoidance has been demonstrated in adult insects (reviewed in Hill, 2008). There is also evidence to suggest that the pupal stage uses vibration. The extra-large larvae of Japanese rhinoceros beetles (*Trypoxylus dichotoma*) beat the prothorax against the pupal wall (see the supplementary videos at bit.ly/3yJOY9y). The vibrations are produced in response to approaching larvae, the first evidence for communication between larvae and pupae (Kojima et al., 2012).

Eavesdropping on Soil Vibrations
Who else might be listening to such sounds and vibrations in the soil? Early evidence suggests that the plant roots themselves are eavesdropping (Thode, 2019). For example, pea shoots (*Pisum sativum*), given a "choice" between growing through two tubes to reach nutrients, have been found to grow toward the vibrations/sounds of water alone, even when physical moisture was absent (Gagliano et al., 2017). Although receptor mechanisms have not yet been identified, it is hypothesized that the roots may be detecting vibrations in a frequency-selective way.

Mammals such as the striped skunk (*Mephitis mephitis*) and North American raccoon (*Procyon lotor*) may also be

tapping into vibratory cues. These species dig and over-turn the turf when seeking their below-ground prey (white grubs, Scarabaeidae larvae). Digging is highly localized at grub hotspots. Given the sound production of white grubs discussed previously in **Vibrations Above and Below the Lawn**, vibroacoustic cues may play a role in this foraging activity, an area that our research is currently exploring. Vibrational detection seems particularly likely in skunks, given the stomping behavior performed when threatened (see youtu.be/HLYD5BdNd90) (Crabbe, 1948).

Humans are the other listeners. Distributions of unseen root-feeding pests (white grubs) can be mapped using sounds and vibrations (**Figure 3**) (Zhang et al., 2003). Soil probes, such as microphones or accelerometers pushed into the soil, record the short, pulsed signals below (Brandhorst-Hubbard et al., 2009).

This approach would have great benefits to pest managers. Without manually sampling the soil, grub hotspots

cannot be located accurately, and managers therefore turn to blanket pesticide treatments instead. Vibro-acoustics provides a novel alternative monitoring strategy. Presently, processing these signals requires a combination of manual listening and automated computer algorithms. There is still much to be learned to enable us to distinguish pest from nonpest and between species. This promising methodology is more common in other substrates such as tree trunks, plants, and crops such as fruit, grain, and timber (Mankin et al., 2011).

From the Lawn to the Seashore

On the seashore (**Figure 4**), crabs use vibrations (Popper et al., 2001). The family Ocypodidae (fiddler and ghost crabs) "rap," "rasp," "drum," and "honk" during complex courtship and territorial displays, producing signals in the range of 0.3 to 3 kHz. Male crabs beat the ground either inside or in front of their sandy burrows, signaling to other males and advertising their attractiveness to females. Other crustaceans, such as marsh and mangrove

Figure 3. Current practices for locating white grubs in turfgrass involve manually hand sampling, as illustrated in the photograph of Louise Roberts and Kyle Wickings (A). Vibroacoustic sampling with soil probes, here shown at Green-Wood Cemetery, Brooklyn, New York (see green-wood.com), offers a promising noninvasive monitoring method for characterizing the distribution and abundance of pests in these areas (B and D). Exemplar oscillogram (amplitude units are provided in terms of relative Raven software units; top) and spectrograms of multiple pulses (bottom), hypothesized to be biologically produced, recorded by Louise Roberts and Kyle Wickings (C). Photos by Ramom V. Pereira (A), used with permission, and Louise Roberts (B and D).
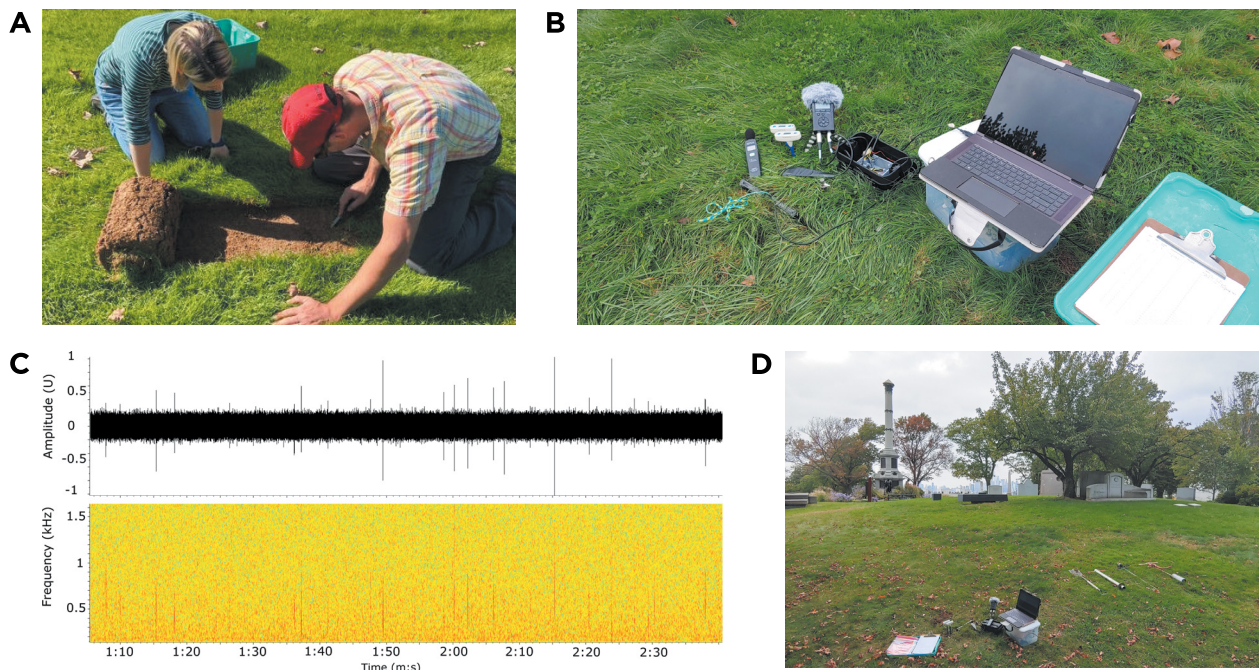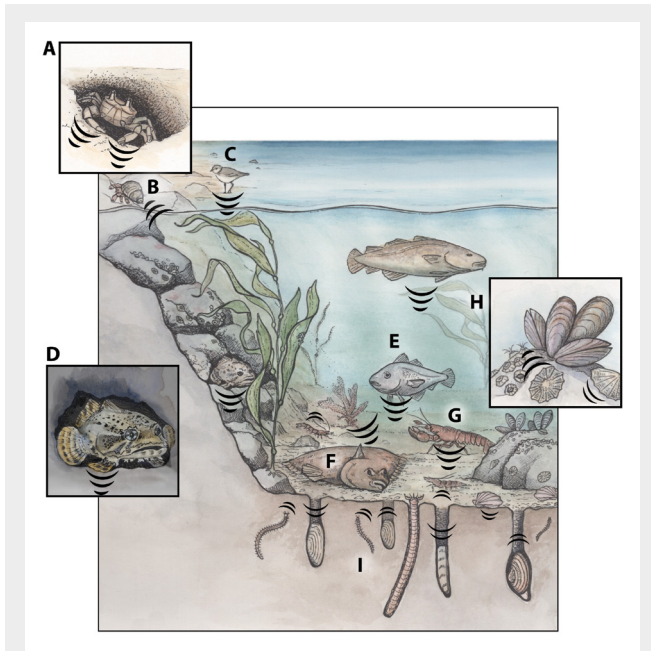
Figure 4. *Cross section of an East Coast aquatic system, with vibrational users and producers. Ghost crabs produce vibrations on the shore (A). Hermit crabs are sensitive to vibrations (B). Wading birds, such as plovers, paddle on the shore to capture prey (C). Oyster toadfish produce loud waterborne calls, hypothesized to translate into the substrate (D). Demersal fish such as haddock and cod also produce waterborne sounds close to the sediment (E). Benthic fish and invertebrates, such as flatfish (F) and brown shrimp (G), plus sessile organisms such as barnacles, limpets and mussels (H) are responsive to vibrations. Burrowing bivalves have known sensitivities to vibrations (I). Paintings by Ann Sanderson, © 2022, Ann Sanderson, all rights reserved; see www.annsciart.com.*

(Sesarmidae and Grapsidae, respectively) crabs produce similar sounds, which may also transmit vibrations into the ground. These signals consist of leg movements such as "stamping," and "tapping." The signals are produced during burrow defense, reproductive contests, and even as postfight victory displays (Goh et al., 209).

Terrestrial hermit crabs (*Coenobita* spp.) that reside on the tropical seashore also produce vibrations. These crabs carry an empty snail shell, a transportable home, to protect their delicate bodies. On occasion, shell fights break out, with attackers attempting to wrench other crabs from their homes. In this context, a repetitive "chirping" sound can be heard, which sounds a little like a frog (0.4- to 11.2-kHz peak) (**Figure 5**) (Roberts, 2021) See **Multimedia File 1** at

acousticstoday.org/robertsmedia that shows a shell contest, with audible "chirps" throughout.

The vibration varies according to the architecture of the shell itself. In this way, by swapping their shells regularly, the hermit crabs may also be varying the information they are "communicating" to other crabs. During earlier stages of these shell contests, defending crabs "shake" in their shell when touched by attacking crabs. Tests have shown that the amplitude of the shaking vibration defines the response of the attacker. Greater amplitude vibrations were more likely to deter attacks, with crabs spending less time assessing shells and being less likely to flip the shell over to examine it (Roberts and Laidre, 2019). The responses of *Coenobita compressus* to shell shaking can be seen in videos at bit.ly/3N0mPA3.

Vibrational prey capture, discussed in **Vibrations Above and Below the Lawn**, is also observed on the seashore, here exhibited in birds of the intertidal area. The ringed plover (*Charadrius hiaticula*) stands on one leg vibrating the toes rapidly against the surface of the sediment (see youtu.be/jq9_k75foS8). This behavior triggers prey surfacing (lugworms, polychaetes, nematodes, and small crustacea), initiated to increase foraging success (Pienkowski, 1983). In laboratory experiments with a mechanical bird foot (!), researchers found that vibration triggers movement of similar prey species, making prey more visible (Osborne, 1982). Paddling, trembling, and leg shaking are particularly prevalent in plovers and other wading birds as well as gulls (family Laridae; see youtu.be/9yr4ZZUH-YQ).

Aquatic Biotremology

Aquatic biotremology, a research area in its infancy compared with terrestrial studies, encompasses every substrate type below the waves, from soft sediments, hard substrata, artificial structures to algal fronds, and perhaps even sessile animals themselves. Studies should include animals living on the surface of the seabed, such as flatfishes, those burrowing down in the sediment, such as mollusks, or even animals that lay eggs on surfaces, such as squid.

Although biotremologists do not yet know much about vibration production in this system, vibration sensitivity is widespread in aquatic organisms. Mollusks, worms (annelids, nematodes, and polychaetes), coelenterates
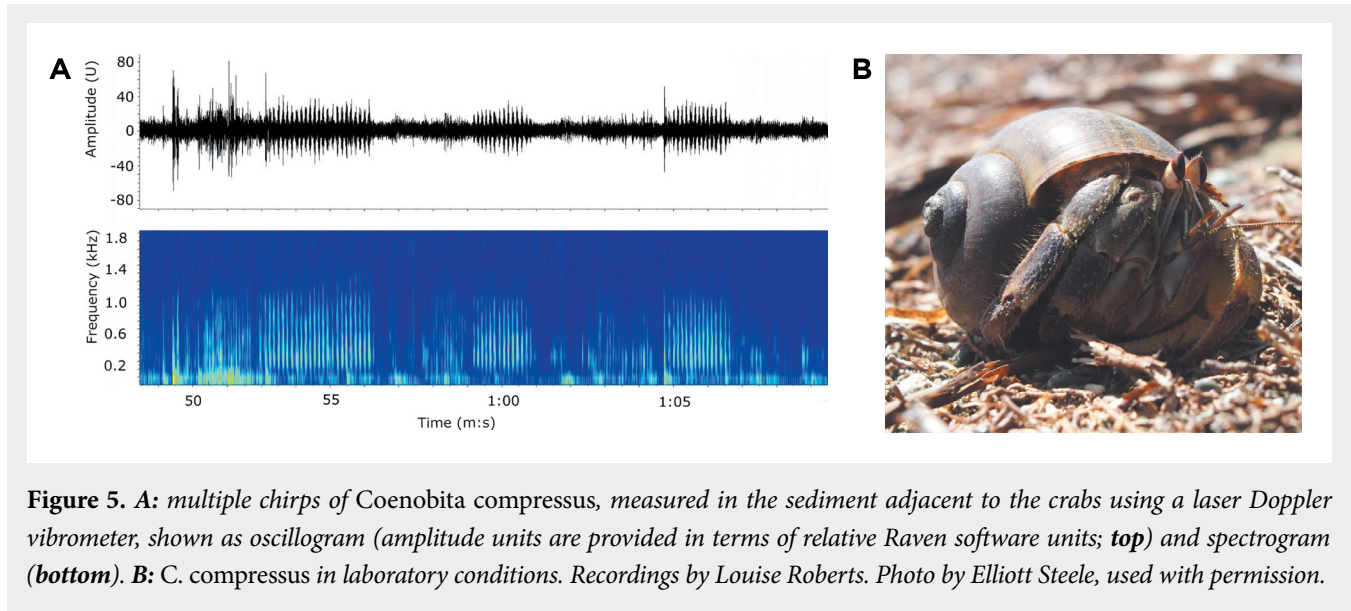
Figure 5. A: *multiple chirps of* Coenobita compressus, *measured in the sediment adjacent to the crabs using a laser Doppler vibrometer, shown as oscillogram (amplitude units are provided in terms of relative Raven software units;* **top***) and spectrogram* (**bottom***).* **B:** C. compressus *in laboratory conditions. Recordings by Louise Roberts. Photo by Elliott Steele, used with permission.*

(anemones, polyps, and comb jellies), arthropods (crustaceans) (reviewed in Budelmann, 1992), and even some sediment-dwelling chordates (e.g., flatfishes) can detect vibrations (Berghahn et al., 1995). Much like a human hearing test, where a sound is played and a response is given by the listener, the vibrational sensitivity of animals can be assessed by monitoring repetitive behavioral responses. Blue mussels (*Mytilus edulis*) fully or partially close their shells after exposure to sediment vibrations, allowing their frequency responsiveness to be quantified (Roberts et al., 2015). Similarly, the response of the hermit crab (*Pagurus bernhardus*) to vibration is predictable enough that monitoring movement changes can inform us of their sensitivities (Roberts et al., 2016).

We have animals sensing vibrations below the waves, but who is producing the vibrations? Certainly, animals moving in, on, and around surfaces will be producing incidental vibrations, but others may be actively signaling, too.

Although speculative at present, aquatic biotremologists may hypothesize that fishes that sit on the seabed could create vibrations when producing waterborne sound. The three-spined toadfish (*Batrachomoeus trispinosus*) produces "hoot" and grunting sounds during mating to attract females to their dens (Rice and Bass, 2009). Similarly, members of the sediment-dwelling Sciaenidae (croakers and drums; see sounds at bit.ly/3GuUi35) are active sound producers (Ramcharitar et al., 2006). Given that these fishes are sitting on or near the bottom when

signaling, it seems highly likely that there is a substrate-borne component to these signals.

Mantis shrimps (*Hemisquilla californiensis*) "rumble" and "rattle" from within their burrows in response to approaching objects (see youtu.be/F2yLsXL74XE). These low-frequency signals also likely propagate through the substrate (Patek and Caldwell, 2006). When vibrational sensors become more standard in aquatic research, it seems probable that vibrational production will be found to be widespread.

How might understanding such vibrations be useful? Much in the same way that bioacousticians can monitor sounds to measure species diversity and abundances, biotremologists can monitor vibrations for the same purpose. This would be particularly useful given that many bottom-dwelling organisms are cryptic in much the same way as soil dwellers. Understanding natural aquatic vibroscapes is also likely crucial to our understanding of the potential impacts of human-made vibrations, as discussed in **Humans Are Shaking the Earth**.

Humans Are Shaking the Earth

The term "noise" is typically used for unwanted sounds that we, as humans, can perceive. Some noises such as lawn mower engines in suburbia or the rumble of traffic from the city center may spring immediately into mind (see Slabbekoorn, 2018). In the water, this noise would include the sound of boat engines for a swimmer or the

sound of bubbles when diving down to the ocean depths. Noises propagating in these media (air and water) have gained biologists' research attention largely because our ears are able to detect these sounds.

However, noise also exists in the context discussed here, as surface vibrations (Raboin, 2021). In this medium, human contributions overlap in frequency and amplitude with the natural vibroscape. Here, the term "vibrational noise" is used, defined as being any anthropogenically produced vibration. For clarity, we use "acoustic noise" to refer to the air- or waterborne equivalent.

Many of the sources associated with acoustic noise are in direct contact with a solid (e.g., the ground but also other surfaces such as plant stems or tree trunks) and thus create vibrational noise (Roberts and Howard, 2022). In both terrestrial and aquatic systems, human activities relating to farming, fishing, construction, energy development/production, mineral exploitation, and transportation all have parts directly in contact with the earth. Foundational structures of wind turbines, for example, directly contact the ground in both on- and offshore systems, making vibration a concern in addition to sound (Popper et al., 2022). Even those activities not in contact with the ground may produce sounds that indirectly translate into the substrate, such as seismic surveys (Hawkins et al., 2021).

As with acoustic noise, vibrational noise may be impulsive or continuous or broadband or tonal. Sources may be mobile or stationary or single point or multiple sourced. Noise may radiate through the substrate, air, water, or perhaps all three. Yet the vibroscape is by no means quiet naturally. Other vibrational contributions include plants and animals (incidental or active) and abiotic sources such as air, sediment, and water movement. Additional vibrational noise may mask cues and signals, distract, and elicit changes at behavioral, physiological, and physical levels in much the same way as acoustic noise (Popper and Hawkins, 2019).

Despite the prevalence of vibrational noise, research relating to potential impacts is sparse, with less than 25 species tested across all environments (reviewed in Roberts and Howard, 2022). The available aquatic data relate mostly to bivalve mollusks and crustacea. Here, behavioral changes have been observed, which include "flinching," burial, and siphonal retraction in bivalves

such as cockles (*Cerastoderma edule*) and scallops (*Pecten fumatus*). In crustaceans, locomotory bursts, impaired feeding, changes in environmental information use, and modified antipredator responses have been observed.

Increased mortality, compromised homeostasis (scallops), and physical damage (rock lobsters; *Jasus edwardsii*) have been demonstrated after exposure to vibroacoustic sources. On land, vibrational noise elicts stress in mice and in farm animals such as cattle, pigs, and chickens. For insects, noise has been shown to distract from producing signals and detecting cues, impact activity budgets, and affect reproduction (parental care, offspring numbers, pair formation).

Below the ground, a recent study found that earthworm abundance decreased with increasing wind farm turbine vibrations but that the vibrational noise did not impact small invertebrates (Velilla et al., 2021).

Taken together, these studies suggested that vibrational noise may impact animal fitness in similar ways to acoustic noise. (Kight and Swaddle, 2011). However, the small number of studies in this area to date mean that it is difficult to draw firm conclusions so far, and there is much to be learned.

The Future for Applied Biotremology
Based on the ubiquity of animal use of vibrational stimuli and that many anthropogenic sources produce such signals, there is clearly an urgent need for experiments evaluating the potential impacts of vibrational noise. How might this be done? The task can be approached with methodologies already found in bioacoustics. Animals may be exposed to actual or replicated anthropogenic sources in carefully controlled exposure or playback studies. However, unlike airborne and waterborne studies, vibrational studies have an additional layer of complexity relating to how vibrations pass through different substrates (Hill, 2008). Acting as a selective frequency filter, the substrate impacts the playback signal, and thus a prefiltration of playback signals is required. Nevertheless, biotremologists use a range of techniques to elicit vibrations in substrates. Piezo actuators may shake plant stems, electromagnetic shakers may tremble a table, thumpers (as in *Dune*) may hit the ground, and ruggedized tactile speakers can shake the ground. If all else fails, hammering a stake into the ground can be sufficient!

But how to study biological vibrations in a world dominated by anthropogenic sources? In the laboratory, air tables, dampeners, gaskets, and suspension reduce external vibrations but typically must be coupled with sound reduction to avoid indirect transmission. It seems then that the rite of passage of a biotremologist is to work in a bunker of one sort or another, in an out-of-the-way campus location, or in a forgotten basement corner.

Regardless of the study location, any research that improves our understanding of vibrational noise is likely to be as critical as the acoustic equivalent, given the prevalence of vibrations used in the animal kingdom for communication. Just a teaser of biotremology has been provided here. The reader might explore this area further in the overview book by Hill (2008).

Acknowledgments

References

Appel, H. M., and Cocroft, R. B. (2014). Plants respond to leaf vibrations caused by insect herbivore chewing. *Oecologica* 175, 1257-1266. https://doi.org/10.1007/s00442-014-2995-6.

Berghahn, R., Wiese, K., and Ludemann, K. (1995). Physical and physiological aspects of gear efficiency in North Sea brown shrimp fisheries. *Helgoländer Meeresuntersuchungen* 49(1-4), 507-518. https://doi.org/10.1007/BF02368378.

Brandhorst-Hubbard, J. L., Flanders, K. L., Mankin, R. W., Guertal, E. A., and Crocker, R. L. (2009). Mapping of soil insect infestations sampled by excavation and acoustic methods. *Journal of Economic Entomology* 94(6), 1452-1458. https://doi.org/10.1603/0022-0493-94.6.1452.

Brownell, P., and Farley, R. D. (1979). Prey-localizing behaviour of the nocturnal desert scorpion, *Paruroctonus mesaensis*: Orientation to substrate vibrations. *Animal Behaviour* 27, 185-193. https://doi.org/10.1016/0003-3472(79)90138-6.

Budelmann, B. U. (1992). Hearing in non-arthropod invertebrates. In Webster, D. B., Fay, R. R., and Popper, A. N. (Eds), *The Evolutionary Biology of Hearing*. Springer-Verlag, New York, NY, pp. 141-155. https://doi.org/10.1007/978-1-4612-2784-7_10.

Catania, K. C. (2008). Worm grunting, fiddling, and charming—Humans unknowingly mimic a predator to harvest bait. *PLoS ONE* 3(10), e3472. https://doi.org/10.1371/journal.pone.0003472.

Crabbe, W. N. (1948). The ecology and management of the prairie spotted skunk in Iowa. *Ecological Monograph* 18(2), 201-232. https://doi.org/10.2307/1948639.

Darwin, C. (1882). *The Formation of Vegetable Mould Through the Action of Worms: With Observations on Their Habits.* Cambridge University Press, Cambridge, UK.

Endler, J. A. (2014). Foreword: The emerging field of tremology. In Cocroft, R., Gogala, M., Hill, P. S. M., and Wessel, A. (Eds.), *Studying Vibrational Communication.* Springer, Berlin Heidelberg, Germany, pp. 7-10. https://doi.org/10.1007/978-3-662-43607-3.

Gagliano, M., Grimonprez, M., Depczynski, M., and Renton, M. (2017). Tuned in: Plant roots use sound to locate water. *Oecologia* 184(1), 151-160. https://doi.org/10.1007/s00442-017-3862-z.

Goh, S. L. R., Vishnu, H., and Ng, N. K. (2019). The sounds of fighting: Contests between violet vinegar crabs, *Episesarma versicolor* (Tweedie, 1940) (Decapoda: Brachyura: Sesarmidae), are resolved through acoustic communication. *Journal of Crustacean Biology* 39(4), 331-341. https://doi.org/10.1093/jcbiol/ruz023.

Görres, C. M., and Chesmore, D. (2019). Active sound production of scarab beetle larvae opens up new possibilities for species-specific pest monitoring in soils. *Scientific Reports* 9, 10115. https://doi.org/10.1038/s41598-019-46121-y.

Grafe, T. (2008). Toe waving in the brown marsh frog *Rana baramica*: Pedal luring to attract prey? *Scientia Bruneiana* 9, 3-5.

Gridi-Papp, M., and Narins, P. M. (2010). Seismic detection and communication in amphibians. *Transworld Research Network* 661(2), 69-84.

Harvey, D. J., Hawes, C. J., Gange, A. C., Finch, P., Chesmore, D., and Farr, I. A. N. (2011). Development of non-invasive monitoring methods for larvae and adults of the stag beetle, *Lucanus cervus*. *Insect Conservation and Diversity* 4(1), 4-14.

Hawkins, A. D., Hazelwood, R. A., Popper, A. N., and Macey, P. C. (2021). Substrate vibrations and their potential effects upon fishes and invertebrates. *The Journal of the Acoustical Society of America* 149(4), 2782-2790. https://doi.org/10.1121/10.0004773.

Hill, P. S. M. (2008). *Vibrational Communication in Animals.* Harvard University Press, Cambridge MA.

Hill, P. S. M., and Wessel, A. (2016). Primer: Biotremology. *Current Biology* 26, R181-R191. https://doi.org/10.1016/j.cub.2016.01.054.

Kaufman, J. H. (1989). Stomping for earthworms by wood turtles, *Clemmys insculpta*: A newly discovered foraging technique. *Copeia* 4, 1001-1004. https://doi.org/10.2307/1445299.

Kight, C. R., and Swaddle, J. P. (2011). How and why environmental noise impacts animals: An integrative, mechanistic review. *Ecology Letters* 14(10), 1052-1061. https://doi.org/10.1111/j.1461-0248.2011.01664.x.

Kojima, W., Takanashi, T., and Ishikawa, Y. (2012).Vibratory communication in the soil: Pupal signals deter larval intrusion in a group-living beetle *Trypoxylus dichotoma*. *Behavioral Ecology and Sociobiology* 66(2), 171-179. https://doi.org/10.1007/s00265-011-1264-5.

Laverack, M. S. (1960). Tactile and chemical perception in earthworms—I. Responses to touch, sodium chloride, quinine and sugars. *Comparative Biochemistry and Physiology* 1(2), 155-156. https://doi.org/10.1016/0010-406x(60)90047-5.

Mankin, R. W., Hagstrum, D. W., Smith, A. L., Roda, A. L., and Kairo, M.T. (2011). Perspective and promise: A century of insect acoustic detection and monitoring. *American Entomologist* 57(1), 30-44. https://doi.org/10.1093/ae/57.1.30.

Mason, P., and Narins, P. M. (2011). Seismic signal use by fossorial mammals. *American Zoologist* 41(5), 1171-1184. https://doi.org/10.1093/icb/41.5.1171.

Mitra, O., Callaham, M. A., Smith, M. L., and Yack, J. E. (2009). Grunting for worms: Seismic vibrations cause *Diplocardia* earthworms to emerge from the soil. *Biology Letters* 5(1), 16-19. https://doi.org/10.1098/rsbl.2008.0456.

O'Connell-Rodwell, C. E., Hart, L. A., and Arnason, B. T. (2001). Exploring the potential use of seismic waves as a communication channel by elephants and other large mammals. *American Zoologist* 41(5), 1157-1170. https://doi.org/10.1093/icb/41.5.1157.

Osborne, B. C. (1982). Foot-trembling and feeding behaviour in the ringed plover *Charadrius hiaticula*. *Bird Study* 29(3), 209-212. https://doi.org/10.1080/00063658209476759.

Patek, S. N., and Caldwell, R. L. (2006). The stomatopod rumble: Low frequency sound production in *Hemisquilla californiensis*. *Marine and Freshwater Behaviour and Physiology* 39(2), 11-99. https://doi.org/10.1080/10236240600563289.

Pienkowski, M. W. (1983). Surface activity of some intertidal invertebrates in relation to temperature and the foraging behaviour of their shorebird predators. *Marine Ecology Progress Series* 11, 141-150.

Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napoletano, B. M., Gage, S. H. and Pieretti, N. (2011). Soundscape ecology: The science of sound in the landscape. *BioScience* 61(3), 203-216. https://doi.org/10.1525/bio.2011.61.3.6.

Popper, A. N., and Hawkins, A. D. (2019). An overview of fish bioacoustics and the impacts of anthropogenic sounds on fishes. *Journal of Fish Biology* 94(5), 692-713. https://doi.org/10.1111/jfb.13948.

Popper, A. N., Hice-Dunton, L., Jenkins, E., Higgs, D. M., Krebs, J., Mooney, A., Rice, A., Roberts, L., Thomsen, F., Vigness-Raposa, K., Williams, K. A., and Zeddies, D. (2022). Offshore wind energy development: Research priorities for sound and vibration effects on fishes and aquatic invertebrates. *The Journal of the Acoustical Society of America* 151, 205-215. https://doi.org/10.1121/10.0009237.

Popper, A. N., Salmon, M., and Horch, K. (2001). Acoustic detection and communication by decapod crustaceans. *Journal of Comparative Physiology* A 187(2), 83-89. https://doi.org/10.1007/s003590100184.

Raboin, M. (2021). Inaudible noise pollution of the invertebrate world, *Acoustics Today* 17(2), 32-39. https://doi.org/10.1121/at.2021.17.2.32.

Ramcharitar, J., Gannon, D. P., and Popper, A. N. (2006). Bioacoustics of fishes of the family Sciaenidae (croakers and drums). *Transactions of the American Fisheries Society* 135(5), 1409-1431. https://doi.org/10.1577/t05-207.1.

Rice, A. N., and Bass, A. H. (2009). Novel vocal repertoire and paired swimbladders of the three-spined toadfish, *Batrachomoeus trispinosus*: Insights into the diversity of the Batrachoididae. *Journal of Experimental Biology* 212(9), 1377-1391. https://doi.org/10.1242/jeb.028506.

Roberts, L. (2021). Substrate-borne vibration and sound production by the land hermit crab *Coenobita compressus* during social interactions. *The Journal of the Acoustical Society of America* 149(5), 3261-3272. https://doi.org/10.1121/10.0004988.

Roberts, L., and Elliott, M. (2017). Good or bad vibrations? Impacts of anthropogenic vibration on the marine epibenthos. *Science of the Total Environment* 595, 255-268. https://doi.org/10.1016/j.scitotenv.2017.03.117.

Roberts, L., and Howard, D. (2022). Substrate-borne vibrational noise in the anthropocene: From land to sea. In Hill, P. S. M., Mazzoni, V., Stritih-Peljhan, N., Virant-Doberlet, M., and Wessel, A. (Eds.), *Biotremology: Physiology, Ecology, and Evolution*. Springer, Cham, Switzerland, pp. 123-155. https://doi.org/10.1007/978-3-030-97419-0.

Roberts, L., and Laidre, M. E. (2019). Get off my back: Vibrational assessment of homeowner strength. *Biology Letters* 15(4), 20180819. https://doi.org/10.1098/rsbl.2018.0819.

Roberts, L., Cheesman, S., Breithaupt, T., and Elliott, M. (2015). Sensitivity of the mussel *Mytilus edulis* to substrate-borne vibration in relation to anthropogenically generated noise. *Marine Ecology Progress Series* 538, 185-195. https://doi.org/10.3354/meps11468.

Roberts, L., Cheesman, S., Elliott, M., and Breithaupt, T. (2016). Sensitivity of *Pagurus bernhardus* (L.) to substrate-borne vibration and anthropogenic noise. *Journal of Experimental Marine Biology and Ecology* 474, 185-194. https://doi.org/10.1016/j.jembe.2015.09.014.

Slabbekoorn, H. (2018). Soundscape ecology of the anthropocene. *Acoustics Today* 14(1), 42-49.

Šturm, R., Rexhepi, B., López Díez, J. J., Blejec, A., Polajnar, J., Sueur, J., and Virant-Doberlet, M. (2021). Hay meadow vibroscape and interactions within insect vibrational community. *iScience* 24 (9), 103070. https://doi.org/10.1016/j.isci.2021.103070.

Thode, A. M. (2019). Bearing fruit: Plant bioacoustics is blossoming. *Acoustics Today* 15(4), 48-54. https://doi.org/10.1121/AT.2019.15.4.47.

Velilla, E., Collinson, E., Bellato, L., Berg, M. P., and Halfwerk, W. (2021). Vibrational noise from wind energy-turbines negatively impacts earthworm abundance. *Oikos* 130(6), 844-849. https://doi.org/10.1111/oik.08166.

Warkentin, K. M. (2005). How do embryos assess risk? Vibrational cues in predator-induced hatching of red-eyed treefrogs. *Animal Behaviour* 70(1), 59-71. http://doi.org/10.1016/j.anbehav.2004.09.019.

Yack, J. E., Smith, M. L., and Weatherhead, P. J. (2001). Caterpillar talk: Acoustically mediated territoriality in larval Lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America* 98(20), 11371-11375. https://doi.org/10.1073/pnas.191378898.

Zhang, M., Crocker, R. L., Mankin, R. W., Flanders, K. L., and Brandhorst-Hubbard, J. L. (2003). Acoustic estimation of infestations and population densities of white grubs (Coleoptera: Scarabaeidae) in turfgrass. *Journal of Economic Entomology* 96(6), 1770-1779. https://doi.org/10.1093/jee/96.6.1770.

About the Authors

Louise Roberts
Lou.Roberts@liverpool.ac.uk

School of Environmental Sciences
Department of Earth, Ocean and Ecological Sciences
Liverpool University
4 Brownlow Street
Liverpool, L69 3GP United Kingdom

Louise Roberts is a lecturer (assistant professor) in marine biology at Liverpool University (Liverpool, United Kingdom), where her research specializes in aquatic biotremology and bioacoustics, centered around human impacts. Prior to Liverpool, Louise was a postdoctoral researcher in the laboratory of Kyle Wickings, Cornell University (Geneva, New York), where her research focused on vibroacoustic monitoring of below-ground invertebrates.

Kyle Wickings
kgw37@cornell.edu

Department of Entomology
Cornell Agritech
Cornell University
630 W North Street
Geneva, New York 14456, USA

Kyle Wickings is an associate professor of entomology at Cornell AgriTech, Cornell University (Geneva, New York), where he works on the ecology and management of soil animals. The Wickings program explores basic ecological questions about soil animal communities and seeks to develop sustainable practices for managing soil-dwelling insect pests in residential and agricultural ecosystems.

Guitar Sounds: From Wood Vibrations to a Mini Power Plant

Jesús Alejandro Torres

Introduction

When a guitar player attends a party, they are often asked to play. A problem, however, is that this is often on a borrowed guitar. If the guitar player normally plays an electric guitar and the borrowed guitar is acoustic (or vice versa), there are real problems because although electric and acoustic guitars appear to be the same instrument to most people, they are really very different instruments. Indeed, there are even differences in how a musician usually plays the guitar. Have you ever seen

Figure 2. *Vibrating string length in guitars is adjusted by pressing the strings against the metallic frets. Labels show how a vibrating string interacts with a different number of terminals depending on the side of its deflection.*

a musician playing an electric guitar seated in a chair? On the other hand, do you remember any guitarist stand while playing an acoustic guitar to perform a classical song? (See **Multimedia File 1** and **Multimedia File 2** at acousticstoday.org/torresmedia).

All guitars have some interesting details in common as well as many differences. For example, the vibrating string length (measured from the bridge to each fret; see parts of the guitar in **Figure 1**) is fixed because the frets cannot be moved. **Figure 2** shows details of the string length limited by the frets. In contrast, bowed instruments, such as the violin, do not have frets. As a result, violinists can impose any vibrating string length without the distances imposed by the frets positions. Thus, fretless instruments allow more expressiveness in the music because the player has more freedom to jump from one note to another. Still, using frets enables beginners to play the guitar because there is no need to guess the exact position of the finger to obtain a well-tuned note.

Figure 1. *Anatomy of three different types of guitars.* **Left:** *acoustic nylon-string guitar.* **Center:** *electric guitar.* **Right:** *acoustic steel-stringed guitar.*
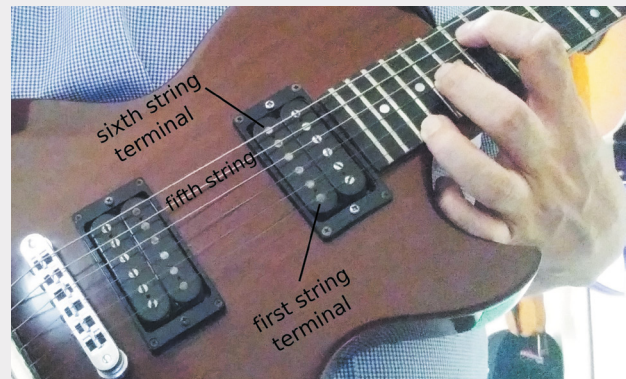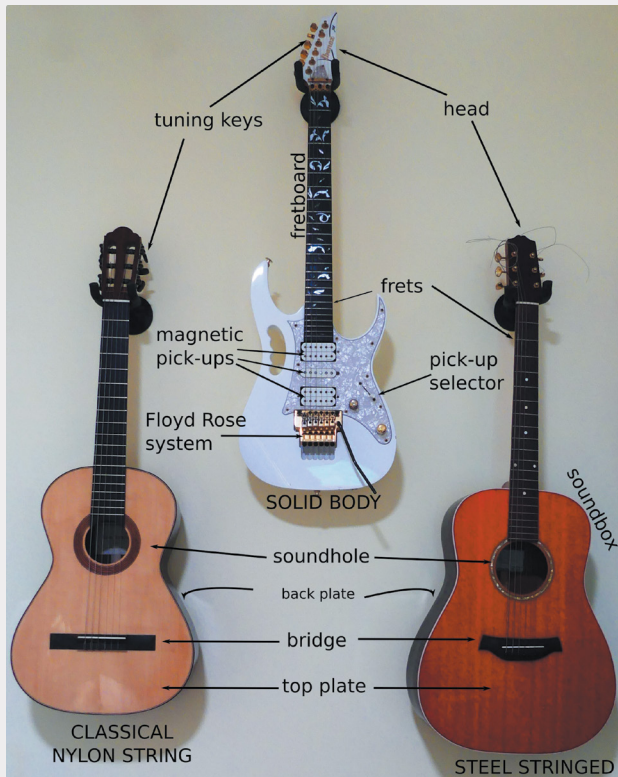
Historical Development of the Guitar

French (2008) points out that anthropologists believe that the oldest forerunners of stringed musical instruments evolved from the stretched strings of bows used by hunters early in the history of civilization. Moreover, it is reasonable to infer that three types of musical instruments evolved from this first instrument: the struck string (e.g., piano), bowed string (e.g., violin), and plucked string, with the guitar arising from the plucked string branch. In fact, instruments with clear similarities to the modern guitar can be seen in a 3,300-year-old stone carving at Alaca Huyuk in Turkey (see tinyurl.com/mvybc2rs). Thus, in a certain sense, the evolution of the design of an instrument like the guitar is probably not very different from the evolution of living things, with slow and somewhat random changes over time and across many cultures.

However, different hypotheses have been proposed to explain the evolution of the guitar. Some people say that it comes from the ancient Greeks, whereas other people think that the guitar comes from early Egyptian instruments. Currently, there is no consensus as to the direct origin of the guitar from any particular ancient instrument. It is clear, however, that the oldest musical scores that can be played on a guitar were written in the sixteenth century, according to Chapman (1993), who also illustrates some ancient instruments like the guitar (also see Bucur, 2016).

During the eighteenth century, stringed instruments such as the Spanish vihuela, a close relative of the guitar, lost popularity because it lacked sound power in comparison with other stringed instruments such as the violin (bowed string) or the piano (struck string). However, the manufacture of plucked instruments was significantly easier than the others, and this helped to increase its popularity.

For the first half of the nineteenth century, in pursuit of a more powerful sound from the acoustic guitar, Antonio de Torres proposed the classical design. We know it now with 65-cm-long strings and a soundbox that included internal bracing. These adaptations provided increased loudness by having a stronger but thinner top and back plates (see **Figure 1**). **Figure 3** shows a remarkable replica of the most iconic guitar made by de Torres, named "La Cumbre." The process of making this replica, including all its artistic details, took the guitar maker Abel García López nine years. **Multimedia File 1** (at acousticstoday.org/torresmedia) shows a

Figure 3. *Left: replica of "La Cumbre" an iconic guitar made by Antonio de Torres dated 1884. Replica made by Abel García López. Artistic details on the soundhole are hown.* **Right:** *the entire guitar. Photograph by José Pita, used with permission.*

performance of this guitar played by Abel García Ayala (see more at youtu.be/1UkoiSm32Rs).

But later, the sound produced by acoustic guitars was again put at a disadvantage when amplifiers using vacuum tubes (or simply "tubes") were developed. By 1931, Rickenbacker and Beauchamp attached a sensor under the strings to capture their vibrations and send the signal to the amplifier, thus producing the first electric guitar sounds. This is discussed in detail by Wilmering et al. (2020), who also share notes about the development of many other audio effects.

The earliest electric guitars were equipped with single coil pickups (see middle pickup on the electric guitar in **Figure 1**). These pickups were sensitive not only to the vibrations of the strings but also to other external magnetic fields, thus producing some noise in the guitar sound. To cut the noise, a second coil was attached to the first one to improve the signal-to-noise ratio. However, the problem with these dual coil pickups (called a "humbucker") was a lack of high-frequency response. Still, humbuckers were widely used in the Gibson Les Paul guitars (see tinyurl.com/2ptuzc5d), and they are still used by guitarists. One of these guitars was used in **Figure 2**. **Multimedia File 2** (at acousticstoday.org/torresmedia) contains a small exercise demonstrating some typical features of playing an electric guitar.

Tuning the Strings

The theory about vibrational behavior of the plucked string can be found in books about fundamentals of

mechanical vibrations and acoustics or even in some about general physics. There are also research papers especially focused on guitar strings. Therefore, instead discussing these topics in detail here, I explain some practical considerations employed by guitar players to obtain different nuances in the notes delivered by the strings to the guitar.

Nominally, guitars have six vibrating strings tuned to E2 (82.4 Hz), A2 (110 Hz), D2 (146 Hz), G2 (196 Hz), B3 (247 Hz), and E4 (330 Hz) to generate the musical notes. Acoustic guitars can be used with nylon strings or steel strings (depending on the music style), but the electric guitars invariably use steel strings for reasons that is discussed in **The Electric Guitar** (see **Figure 1**). Moreover, although most guitars have six strings, a few models add a seventh string, typically tuned to B1 (61.7 Hz), to play notes lower than E2. In addition, players of different musical styles ranging from classical to blues intentionally modify the standard tuning for some pieces by adjusting the tension of the strings using the tuning keys.

According to the pioneering laws published by Mersenne (1636) about the musical sounds of a vibrating string, we know that tuning a string depends on its diameter, material, length, and tension. As mentioned, frets are responsible for limiting the vibrating string length in guitars (**Figure 2**). Therefore, to change a note, one presses the string on a different fret. In addition, gradual variations of the guitar notes can be done without involving changing the fret that is pressed. The most common technique to produce a subtle change in the tuning of a note in the guitar consists of just sliding the finger that is pressing the string over the fret, a technique known as bending. By bending, the tension is greatly increased in comparison with the small increase in the string length, therefore producing an increase in frequency. See Paté et al. (2012) for a brief glossary of techniques for the electric guitarist and the label in **Multimedia File 2** (at acousticstoday.org/torresmedia).

Bending in steel strings can achieve variations up to one tone (see label for bending from D to E in **Multimedia File 2** at acousticstoday.org/torresmedia) or even a little more, whereas the change is less noticeable in nylon strings, although entirely feasible. See Lewis et al. (2014) for a more thorough analysis of the effects of the variations of the tension on a nylon string. In a more

sophisticated technique employed in electric guitars, the bridge can be moved, enabling dramatic changes in the tension of the strings (see the end of **Multimedia File 2** at acousticstoday.org/torresmedia).

Controlling Harmonics in Plucked Strings

If the string is plucked close to the bridge, then its vibrations are rich in harmonics. However, because the high-frequency harmonics use a lot of the vibrating energy of the string, the note fades out quickly. As the string is plucked closer to its midpoint, the harmonic content decreases, but the length of the note increases (see **Figure 4**). It is necessary to consider these characteristics to obtain a realistic simulation of the guitar tone, as demonstrated in Torres and Rendón (2013). Because musicians empirically learn that plucking different parts of the strings creates different sounds, they intentionally vary the position of the plucking hand to control both the harmonic content and the length of the notes.

In addition to the plucking location, the plucking technique and the plucking (attack) angle are both important (Fletcher and Rossing, 2012). Predominantly, the nails (tirando) and the fingertips (apoyando) are used with nylon-string guitars (see **Multimedia File 1** at

Figure 4. *First three seconds of the sound made by plucking the first open string of a steel-stringed guitar. The amplitudes are expressed as decibels relative to full scale. The envelope for the graph of the string plucked near to the bridge (**bottom**) fades out more quickly than when the string is plucked at its midpoint (**top**), but the harmonic content of the **bottom** image is richer.*
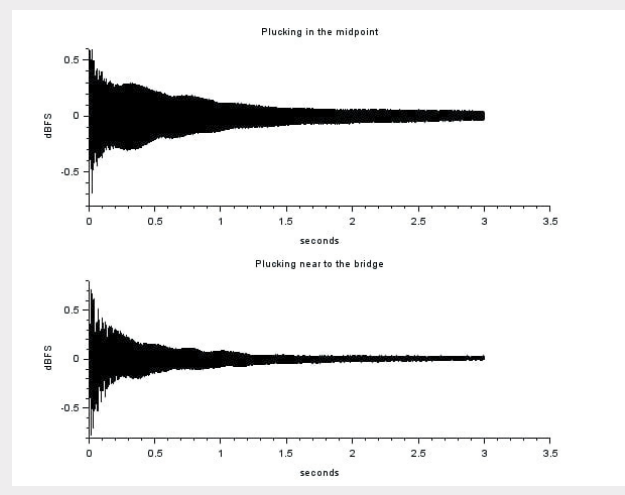
acousticstoday.org/torresmedia). Steel strings are plucked by means of a plectrum, a thin flat piece of plastic or metal held by the fingers (a metal one is shown in the first seconds of **Multimedia File 2** at acousticstoday.org/torresmedia). There are some techniques in which both styles are combined. For example, if the string is plucked with a plectrum and immediately after this the vibrations are damped with the thumb of the same hand, a vibration is produced that only contains high-frequency harmonics. Indeed, in electric guitars amplified with distortion, this technique causes a particularly spectacular effect that is often referred to as artificial harmonics (see labels in **Multimedia File 2** at acousticstoday.org/torresmedia).

Once the string is installed on the guitar and tightened (tensed), neither the diameter nor the material properties seem to remain constant. In fact, it takes a several days for the newly installed strings to reach stable tuning. Thus, experienced guitarists never change strings on the day of an important performance. Also, as time goes by, the strings keep changing because the harmonic content is affected, whereas the damping in the material increases. Thus, over time, the strings gradually wear out because they are used until they break or until their behavior ceases to be the pleasant sound that they should be for the musician. Then, the strings need to be replaced.

The ability of a vibrating string to generate sound waves by itself is poor because its surface in contact with the air, the propagation medium, is too limited. For this reason, once the guitar strings have been plucked, their vibrations must be collected in some way by other systems responsible for amplification of the sound. In acoustic guitars, a small fraction of the vibrating energy of the string is extracted by the bridge to drive the whole soundbox. In electric guitars, string vibrations are collected through magnetic pickups that generate an electrical signal to be sent to an external amplifier.

The Soundbox
There are many design differences, even among acoustic guitars. **Figure 1** compares the different orientation of the tuning keys in the two types of acoustic guitars shown. The soundbox of a steel-stringed guitar is a little bigger and is subject to more tension than that in a nylon-stringed guitar (seen in **Figures 1** and **5**). To prepare the soundbox of a steel-stringed guitar to support more tension, the internal design of the top plate includes an
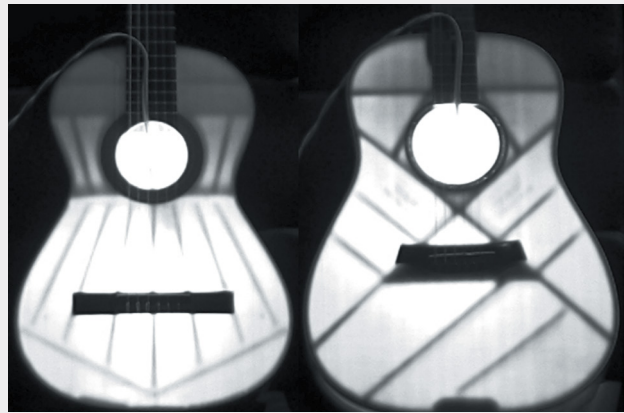
Figure 5. *Internal structures of the acoustic guitars of* **Figure 1**. *The nylon-string guitar shows the classical Antonio de Torres fan bracing design (**left**). The steel-stringed guitar shows X-bracing (**right**). Infrared light from inside was employed to reveal the structure.*

X-bracing instead of the typical fan bracing employed in the classical nylon-stringed guitars.

Figure 5 shows an internal visualization of the two types of acoustic guitars using infrared light (Torres et al., 2010). Although the reinforcements of the top plate provide structural functions, they also radically affect the sound of the instrument because they alter the vibrational behavior of the top plate. The top plate is largely responsible for the sound amplification in the acoustic guitar, as shown in the simulations made by Torres and Torres-Torres (2015). **Multimedia File 4** (at acousticstoday.org/torresmedia) shows a mode shape of a top plate for middle frequencies, with deflections clearly constrained by the fan bracing of the acoustic guitar.

A musical excerpt played on an acoustic guitar made of wood will never sound exactly like the same excerpt played on any other acoustic guitar, even if both instruments have the same design and the same type of strings. But why is the sound of each acoustic guitar unique? It is because the soundbox is responsible for the sound amplification in the acoustic guitars, and the response of each soundbox to the vibrations coming from the strings is unique. To explain more about that, it is necessary to analyze how acoustic guitars work in a little more detail.

To analyze a guitar's performance, it is very useful to study its vibrational behavior in the frequency domain.

For this purpose, the vibrations are measured in terms of motion per unit of input force to the instrument. The typical procedure consists of hitting the bridge with an impact hammer, sensing the velocity caused as a response, and relating both signals (velocity/force) to one another, thus obtaining a transfer function called bridge mobility. A graphic obtained in this way from the "La Cumbre" replica is shown in **Figure 6**, where, with a bit of experience, the contribution of the two main vibratory systems can be easily identified. The first three well-separated peaks are caused by a strong coupling of a Helmholtz-like resonance with the soundbox, and peaks beyond those are mainly due to resonances of the soundbox itself.

The enclosed air of the soundbox supports the amplification of the lowest frequencies of the instrument, mainly by a resonance that resembles that produced when one blows across the hole of an open bottle. Indeed, if the adequate flow of air is directed to the sound hole of an acoustic guitar, one hears the sound of the frequency amplified by this resonance (**Multimedia File 3** at acousticstoday.org/torresmedia). Such behavior is the first one that emerges in the response of acoustic guitars and usually appears at around 100 Hz. The mobility of the "La Cumbre" replica (**Figure 6**) revealed a particularly low frequency for this resonance, matching the fundamental frequency of the E2 note (82.4 Hz). **Multimedia File 1** (at

Figure 6. *Bridge mobility of the "La Cumbre" replica. This is a typical measurement used by scientists to study the vibrational behavior of the instrument.*

acousticstoday.org/torresmedia) contains a piece played in E tonality using the "La Cumbre" replica.

The rest of the frequencies of the sound of an acoustic guitar, in the mid and high range, are radiated by the wood itself. Therefore, almost all the sound of the guitar depends on the modal behavior of the soundbox. This, in turn, depends on both the design of the soundbox and the physical properties of the wood. Thus, the same design employing different materials will result in guitars with different responses, such as different samples of wood, even when they are from the same tree (see Torres and Torres-Martinez, 2015). Moreover, most of the processes involved to produce acoustic guitars of wood are handmade, and there are several designs. As a result, each soundbox for the body of a guitar is unique (Skrodzka et al., 2011).

To explain the relationship between the vibrations of the soundbox and the sound generated, we need to learn more about the modal behavior of the structure. The soundbox of a guitar has numerous resonances whose frequencies have no relationship with one another, which is different from the vibrations in strings that show a harmonic series. In the soundbox, resonances depend on the geometry of the structure, mass, elastic properties, and even variations in humidity (Torres et al., 2014). Nevertheless, the first lowest modes of the acoustic guitars tend to be similar despite the variations of the internal bracing.

Figure 7 is a schematic of responses of a 1965 Ramirez guitar, printed by Graham Caldersmith in 1980 where he labeled some peaks with the corresponding mode shapes. In turn, **Multimedia File 5** (at acousticstoday.org/torresmedia) shows the visualization of the first two modes in an experimental guitar (with strong ribs to isolate the vibrations of its top plate) corresponding to modes (0,0) and (1,0). Both modes are hand drawn over the two biggest peaks of the response plotted in **Figure 7**. The harmonics of notes plucked in the strings, matching the resonant frequencies of the soundbox of each guitar, will be selectively amplified by modes with a high radiation efficiency (Torres and Boullosa, 2011), but harmonics with frequencies far from resonances will be scarcely amplified. More details about this experiment in efficiency are available in Torres and Boullosa (2009).
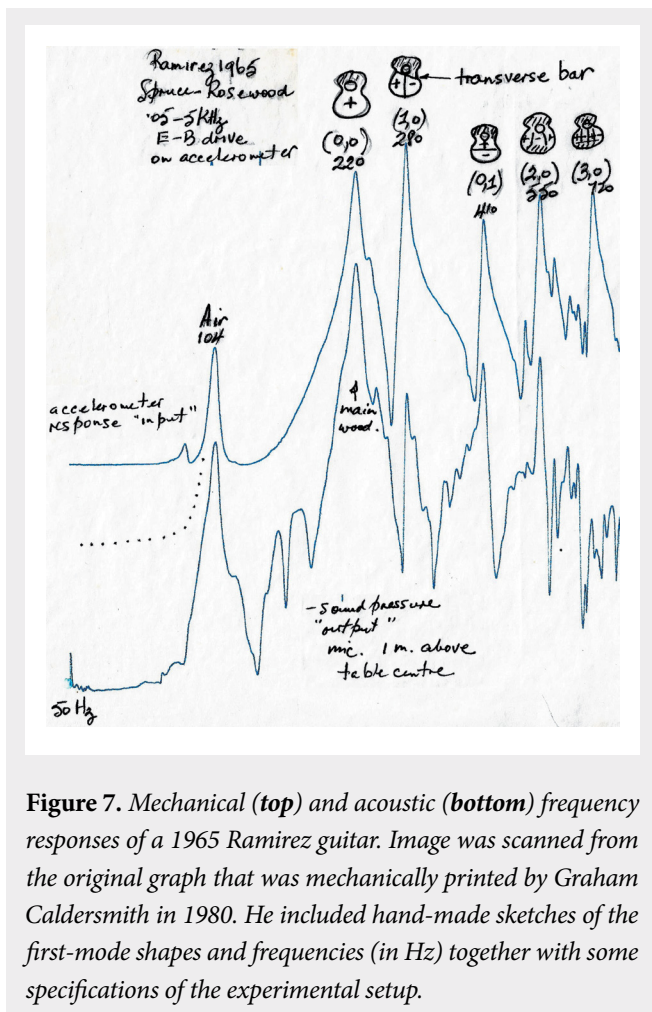
Figure 7. *Mechanical (**top**) and acoustic (**bottom**) frequency responses of a 1965 Ramirez guitar. Image was scanned from the original graph that was mechanically printed by Graham Caldersmith in 1980. He included hand-made sketches of the first-mode shapes and frequencies (in Hz) together with some specifications of the experimental setup.*

The Electric Guitar

A general review of the functionality of the electric guitar is available in books about musical instruments (e.g., Fletcher and Rossing, 2012; Bucur, 2016). In most electric guitars, only the vibrations of the strings are sensed by a magnetic pickup. Because there is no need for vibrations of the guitar body, a solid body substitutes for the soundbox. Moreover, because the energy of the strings' vibrations remains in the strings, the decay of the notes in electric guitars is slower than in acoustic guitars (Paté et al., 2014). Thus, electric guitar players usually damp the vibrations of the strings by gently pressing them on the bridge while they are being plucked using the palm-muting technique (hear the sound obtained using this technique in **Multimedia File 2** at underlinedacousticstoday.org/torresmedia). On the other hand, the remarkable sensitivity to small vibrations in the strings, together with the external acoustic energy provided by its amplified loudspeaker, can achieve an infinite sustaining of the notes because of feedback.

Understanding the operating principles of electric guitars require a good deal of physics and mathematics that are discussed in Horton and Moore (2009) and are not considered here. It is important to understand that the basic mechanism of these guitars involved an electromagnetic pickup consisting of a very thin conductive wire coiled to form a spiral around a permanent magnet. The wire is glued to the magnet so that there is no relative motion between them. Because the steel guitar strings are very close to the pickup, they are inside its magnetic field, resulting in a contactless interaction between the vibrating strings and the magnetic field of the pickup system. Therefore, if nylon strings are plucked in an electric guitar, the instrument simply does not work.

Strings disturbing a magnetic pick-up generate signals completely unrelated to the waveform amplified by the body of an acoustic guitar. Indeed, the notes generated by an electric guitar are not like those of any other musical instrument. Moreover, the waveform that is generated by each terminal of the same guitar pickup tends to be different for each string.

To illustrate why, let us analyze the vibrations of the fifth string of a standard electric guitar and the interaction with the terminals of the pickup (see labels in **Figure 2**). Imagine that the oscillations occur in a plane that is parallel to the top of the pickup. In addition to the interaction of the string with its corresponding terminal of the pickup, during the deflections toward one side, the string will be closer to the terminal of the sixth string. When the string moves in the opposite direction, it will have interactions with the terminal of the fourth, third, second, and first strings. Consequently, the signal is markedly asymmetrical, with a high harmonic content for each waveform. Also, such asymmetry will be different for each string depending on its position over the pickup. These interactions have been elegantly described by Horton and Moore (2009).

The peculiar signal generated by the pickup is only a part of the typical sound of the electric guitar. Additionally, the performance of the amplifier used to produce the sound waves is extremely important. It is interesting to consider that when the electric guitar was developed, the only way to amplify sound with a loudspeaker was using tubes. This type of amplifier (named a valve-state amplifier) has no linear behavior, so its output signal is not a

larger version of the input signal. Therefore, the waveform is distorted. However, because there was no other option to hear the sounds of an electric guitar, valve-state amplifiers were important during the definition of the original sound of this new instrument.

With the development of solid-state amplification some years later, the signal could be amplified, avoiding distortion and the distinctive sound obtained through the tubes. It is interesting, however, that guitar players never really accepted the distortion-free sound of solid-state amplifiers, and so they have kept alive the manufacture of valve-state amplifiers to the present day in pursuit of the original sound (with the inherent distortion through the tubes). Valve-state amplifiers for electric guitars, it is estimated, consume as many as three out of four of the world's production of audio tubes (Barbour, 1998).

Tube amplification causes distortion under two conditions: turning up the volume of the amplifier or sending a powerful signal from the guitar (see **Multimedia File 2** at acousticstoday.org/torresmedia). Usually, electric guitar players try to find the exact volume on the amplifier where, if they subtly play the strings, a clear sound without distortion is obtained, whereas an aggressive attack causes overdrive. A very enlightening discussion about this was published by Keeports (2017).

One of the most remarkable additions to the electric guitar was a mechanical system to change the position of the bridge during a performance. It is installed in several models, and every electric guitar player knows it by the name of the inventor: the Floyd Rose system (Rose 1979) (**Figure 1**). This ingenious mechanism allows dramatic changes in the tension of the strings by moving the bridge and returning it to the original position. The performance of the Floyd Rose system of the electric guitar of **Figure 1** is seen at the end of **Multimedia File 2** (at acousticstoday.org/torresmedia).

Conclusions

If it is assumed that the key role of guitars is producing musical sounds, from what has been established in this article, the conclusion could be that the acoustic guitar and the electric guitar are two almost unrelated instruments. Indeed, the sound of the two instruments is unmistakably different. **Multimedia File 6** (at acousticstoday.org/torresmedia) shows a brief demonstration

comparing the same notes played on the steel-stringed guitar and a single-coil electric guitar.

The quick fade out of the high-frequency harmonics in the sound of the acoustic guitar means that when several notes are played concurrently, there is a clean and well-defined mix of sounds. Because of this characteristic, the acoustic guitar can be considered a polyphonic instrument, meaning that several notes can be played at the same time. Acoustic guitar players can pluck complete chords to provide musical accompaniment, and with more skill, a melody can be played on the same guitar at the same time. Therefore, the support of additional musical instruments (as a piano or a bass) is not required. Most pieces written for the classical guitar are conceived without involving additional instruments.

In contrast, the high-frequency harmonics are present for a longer time in the notes delivered by the electric guitar due to the setup of the magnetic pickup and the nonlinearity of the valve-state amplifier. Although it is possible to obtain clean chords with the adequate setup in electric guitars, the characteristics that became popular are identified when its sound is distorted. Under these conditions, the electric guitar is employed as a monophonic musical instrument, which means that mixing the sound of several notes at the same time tends to be unpleasant or even perceived as noise. However, beautiful melodic lines can be created with the musical accompanying of at least a guitar bass. Occasionally, some nuances or rhythms can also be played using two or a maximum of three notes plucked at the same time, but it is very unusual to simultaneously play the six strings, unlike the case for the acoustic guitar.

Often young guitar students try learning to play both the acoustic guitar and the electric version. But studying each one requires different techniques and a lot of effort. Then, after a certain point, one of the guitar types seduces the musician more than the other type. How does this choice happen? Well, it is hard to explain, but undoubtedly some of the personality of the guitar player will be reflected in the type of guitar selected: acoustic or electric.

Acknowledgments

and to Angela McPherson for sharing several invaluable materials written and measured by Graham Caldersmith.

References

Barbour, E. (1998). The cool sound of tubes. *IEEE Spectrum* 35(8), 24-35.

Bucur, V. (2016). *Handbook of Materials for String Musical Instruments*. Springer, Cham, Switzerland.

Chapman, R. (1993). *The Complete Guitarist*. Dorling Kindersley Publishing, London, UK.

Fletcher, N. H., and Rossing, T. D. (2012). *The Physics of Musical Instruments*. Springer Science and Business Media, New York, NY.

French, R. M. (2008). *Engineering the Guitar: Theory and Practice*. Springer Science and Business Media, New York, NY.

Horton, N. G., and Moore, T. R. (2009). Modeling the magnetic pickup of an electric guitar. *American Journal of Physics* 77(2), 144-150.

Keeports, D. (2017). The warm, rich sound of valve guitar amplifiers. *Physics Education* 52(2), 1-7.

Lewis, W. J., Smith, J. R., Lewis, W. J., and Smith, J. R. (2014). The effect of string tension variation on the perceived pitch of a classical guitar. *Exchanges: The Interdisciplinary Research Journal* 2(1), 53-81.

Mersenne, M. (1636). *Harmonie universelle, contenant la théorie et la pratique de la musique*, S. Cramoisy, Paris, France. Available at https://tinyurl.com/mrjbarmx.

Paté, A., Le Carrou, J. L., and Fabre, B. (2014). Predicting the decay time of solid body electric guitar tones. *The Journal of the Acoustical Society of America* 135(5), 3045-3055.

Paté, A., Navarret, B., Dumoulin, R., Le Carrou, J. L., Fabre, B., and Doutaut, V. (2012). About the electric guitar: A cross-disciplinary context for an acoustical study. *Proceedings of the Acoustics 2012 Nantes Conference,* 11th Congrès Français D'Acoustique: 2012 IOA Annual Meeting, Nantes, France. April 23-27, 2012. Available at https://tinyurl.com/46p6u558. Accessed May 10, 2022.

Rose, F. D. (1979). *Guitar Tremolo Method and Apparatus.* US Patent No. 4,171,661, US Patent and Trademark Office, Washington, DC, October 1979. Available at https://tinyurl.com/ybd3w7df. Accessed May 10, 2022.

Skrodzka, E., Łapa, A., Linde, B. B., and Rosenfeld, E. (2011). Modal parameters of two incomplete and complete guitars differing in the bracing pattern of the soundboard. *The Journal of the Acoustical Society of America* 130(4), 2186-2194.

Torres, J. A., and Boullosa, R. R. (2009). Influence of the bridge on the vibrations of the top plate of a classical guitar. *Applied Acoustics* 70, 1371-1377.

Torres, J. A., and Boullosa, R. R. (2011). Radiation efficiency of a guitar top plate linked with edge or corner modes and intercell cancellation. *The Journal of the Acoustical Society of America* 130(1), 546-556.

Torres, J. A., and Rendón, P. L. (2013). A simple method for synthesizing and producing guitar sounds. *European Journal of Physics* 34(3), 503-510.

Torres, J. A., and Torres-Martinez, R. (2015). Evaluation of guitars and violins made using alternative woods through mobility measurements. *Archives of Acoustics* 40(3), 351-358.

Torres, J. A., and Torres-Torres, D. (2015). Changes in wave propagation in a guitar top plate due to the fan bracing and the bridge. *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería* 31(4), 228-234.

Torres, J. A., de Icaza-Herrera, M., and Castaño, V. M. (2014). Guitar acoustics quality: Shift by humidity variations. *Acta Acustica united with Acustica* 100(3), 537-542.

Torres, J. A., Hernandez, G., Granados, A., and Garcia, A. (2010). Internal visualization of finished guitars and their sound. *Proceedings of Meetings on Acoustics* 11(1), 035005.

Wilmering, T., Moffat, D., Milo, A., and Sandler, M. B. (2020). A history of audio effects. *Applied Sciences* 10(3), 1-27.

About the Author

Jesús Alejandro Torres
jesusalejandrott@yahoo.com.mx

Laboratorio de Acústica
Escuela de Laudería
Instituto Nacional de Bellas Artes
y Literatura
Hidalgo #20 Centro
Querétaro, Querétaro, México,
CP 76000

Jesús Alejandro Torres is a researcher of stringed instruments. All the work for his bachelor's degree in mechanical engineering (2003), master's degree (2006), PhD (2010), and postdoctoral position was focused on experimental testing and computer simulations about the vibroacoustic behavior of the acoustic guitar. Since 2011, Dr. Torres has led the acoustics lab at the Violin Making School of Mexico, where he expanded his interest to study the violin. He currently teaches a two-semester course on acoustics of the violin. He is also an enthusiastic electric guitar player.

Recent Acoustical Society of America Awards and Prizes

Acoustics Today is pleased to announce the names of the new Fellows in the Acoustical Society of America. After the members are approved by the Executive Council of the Society at each semiannual meeting, their names are published in the next issue of *Acoustics Today*. For more information on the accolades, please see acousticalsociety.org/procedures#fellowship.

Congratulations to the following members who were elected Fellows in the Acoustical Society of America, who will formally be recognized at the Fall 2022 Plenary Session.

- **Brian E. Anderson** (Brigham Young University, Provo, Utah) for contributions to applications of acoustic time reversal and acoustics education
- **John J. Galvin III** (House Ear Institute, Los Angeles, California) for research into speech and music perception with cochlear implants

- **Murray S. Korman** (United States Naval Academy, Annapolis, Maryland) for advancing acoustics education through mentoring student acousticians and developing innovative demonstrations
- **Steven M. Lulich** (Indiana University, Bloomington) for studies of speech production using three-dimensional ultrasound imaging
- **Andrew A. Piacsek** (Central Washington University, Ellensburg) for exceptional service to the Society, including leadership in scientific communication of acoustics
- **Lina Reiss** (Oregon Health and Science University, Portland) for studies of combined electric and acoustic hearing
- **Yue Wang** (Simon Fraser University, Burnaby, British Columbia, Canada) for studies of the behavioral and neural mechanisms underlying speech learning and processing

Conversation with a Colleague: Andone Lavery

Andone Lavery
Conversation with a Colleague Editor:
Micheal L. Dent

Meet Andone Lavery

Andone Lavery is a senior scientist in the Department of Applied Ocean Physics and Engineering at the Woods Hole Oceanographic Institution (WHOI; Woods Hole, Massachusetts). She received her degree in mathematics as well as an MMath from the Department of Applied Mathematics and Theoretical Physics at the University of Cambridge (Cambridge, United Kingdom) and her PhD in physics from Cornell University (Ithaca, New York). She completed her postdoctoral training at WHOI. Andone is a Fellow of the Acoustical Society of America (ASA) and has served as an associate editor for *The Journal of the Acoustical Society of America* (*JASA*) and chair of the Acoustical Oceanography Technical Committee. We asked Andone to give us her elevator pitch and then to elaborate on her inspirations, contributions, and hopes for the future.

Give your "elevator speech" about the thrust(s) of your scholarly work over your career.

Fundamentally, the engine that drives research in acoustical oceanography today is the vast variability that characterizes our oceans. This oceanographic variability spans scales of seconds, (e.g., surface waves) days, (e.g., storms) all the way to decades, (e.g., large ocean currents, part of the ocean conveyor belt) and longer. Similarly, the spatial scales associated to these processes span millimeters in phenomena such as capillary waves, bubbles, and turbulence to hundreds of kilometers accompanying the basin scale process, for example, the formation and dynamics of the Gulf Stream. This physical oceanographic variability forces biological variability. For example, plankton patchiness can play an important role in ecosystem function by impacting biomass distribution across trophic levels, thus impacting energy transfer and marine food web dynamics. Recently, spatial scales and timescales associated with anthropogenic impacts have also become important, but their impacts are relatively unknown at this time. For example, marine heat waves and an apparent increase in the number of warm core rings, a type of mesoscale eddy shed from a large ocean current such as the Gulf Stream, have a relatively unknown impact on ecosystem function. All this variability poses a daunting sampling problem, exacerbated by the challenges in accessing remote sites for field work because the ocean is not always a friendly and cooperative colleague. The use of sound to explore the oceans, to understand this variability, and to make the ocean more "transparent" has driven my research throughout my career.

What inspired you to work in this area of scholarship?

I grew up like driftwood by the ocean, in a family of surfers, moving beach break to beach break, across countries and continents, to find the perfect wave, with one of the only things going for me was that I was surprisingly inspired by and luckily relatively good at mathematics, which bought me a ticket to the University of Cambridge. But the ocean did not call to me at that point, and I was not inspired to study and explore the ocean until much later in life.

After graduating from the University of Cambridge, I decided to become an astrophysicist and ended up heading to the Cornell University Physics Department as a graduate student. I was fortunate that most physics graduate students at Cornell University had to teach undergraduate physics, including teaching physics labs to undergraduates. Not only was this the forum where I actually learned physics, but it introduced me to physics lab work, which I was immediately fascinated by and the main reason I ended up sidestepping a career as an astrophysicist. My graduate research focused on understanding nonadiabatic electron transfer processes during reactive ion-surface collisions, specifically looking at charge transfer and memory loss of low-energy ions scattering from single-crystal surfaces.

As is true for many in ocean acoustics, I landed in the ocean acoustics area almost by accident. While a graduate student at Cornell University, studying condensed matter physics, I met a biological oceanographer at a graduate seminar, Chuck Greene, who encouraged me to reach out to colleagues in the Acoustics Laboratory at WHOI. I was fortunate enough to be scientifically adopted by Tim Stanton, who has become my long-standing colleague and still my mentor, and joined a number of his expeditions focused on acoustic mapping of copepods in the Gulf of Maine. Soon after completing my PhD, I joined the Department of Applied Ocean Physics and Engineering, with no formal training or pedigree in oceanography, engineering, or acoustics.

During my time as a postdoc, I focused on my strengths as a trained physicist, delving deeply into the development of physics-based acoustic scattering models for different marine organisms and small-scale physical processes such as oceanic microstructure and double diffusion. While many of my colleagues at WHOI focused on long-range, low-frequency acoustic propagation, I was more driven to understand the interaction of sound with smaller spatial scales in the ocean, propelled by questions in ecosystem acoustics and coastal oceanography. I was funded by the Office of Naval Research Code 32 to develop physics-based acoustic-scattering models for microstructure and double diffusion and to perform laboratory experiments to verify these models. Acoustic techniques have the advantage of being one of the few effective remote-sensing techniques in the ocean, a distinct advantage for large-scale mapping of small-scale

processes needed for understanding how the ocean is effectively stirred and mixed by winds, waves, and tides.

Of all your contributions during your career, which are you most proud of and why?

One of the key driving forces in the area of acoustic scattering for classification and quantification of oceanic "targets," basically any discrete object (e.g., bubble, sediment particle, plankton) or extended process (e.g., stratification, turbulence, fronts) that has an acoustic impedance in the ocean that can scatter sound, is the need to understand the scattering signatures of the targets, that is, a target's individual "fingerprint." I was fortunate enough to enter the field of acoustical oceanography while it was transitioning from active, single-frequency, or narrowband approaches to multifrequency or broadband approaches for classification and quantification. I happened to be in the right place at the right time to lead efforts in the nascent area of broadband field-based acoustical oceanography.

The Shallow Water 2006 Experiment (see [doi.org/10.1121/1.2972156](doi.org/10.1121/1.2972156)), led by Jim Lynch at WHOI, was my first involvement in a large ocean acoustics field effort. According to Jim, enough instruments were deployed to cause the sea level to rise! My own contribution, as an assistant scientist at WHOI at the time, was to deploy a high-frequency broadband acoustic backscattering system to characterize internal waves and zooplankton. I was able to measure the acoustic-scattering signature of microstructure, small-scale fluctuations in temperature and salinity caused by stratified turbulence and spectrally distinguish these signals from those of scattering from zooplankton. I had spent some time working on physics-based acoustic-scattering models of microstructure and was gratified to find that my predictions were in good agreement with the data, especially as there were no free parameters in my model.

I had a growing family at this time in my career, and I actively decided to focus my career on scientific questions that I could address in my own "backyard" without the need to travel extensively to remote field locations. Looking back, if I have one career regret, it was not participating in the GLOBEC Southern Ocean experiments with my colleague Peter Wiebe at that time and the opportunity to study Antarctic ecosystems.

My decision to remain local launched years of research and exploration into estuarine acoustics. I had come to realize that acoustic scattering from microstructure in most open-ocean environments was typically dominated by small-scale fluctuations in temperature. However, it was the exceptions to this "rule" that intrigued me. Salt-finger double diffusion, convective double diffusion, and many estuaries are characterized by strong gradients in salinity, in turn resulting in strong small-scale salinity fluctuations. These salinity fluctuations typically persist down to smaller scales than temperature microstructure, and traditional in situ approaches to measurements of salinity are challenging because they require resolving much smaller scales and additionally require coincident measurements of conductivity and temperature. Resolving both temperature and salinity microstructure using traditional oceanographic technologies at submillimeter scales remains challenging to this day. And yet these are needed to fully resolve the dissipation of rate of salinity variance, an important variable in characterizing turbulent mixing in many environmental flows. High-frequency acoustic techniques allow the salinity variance to be measured remotely at these small scales, and it is sufficiently spectrally distinct from temperature microstructure that it can be distinguished with sufficient broadband signals.

I spent many years mapping salinity variance in the Connecticut River with my colleague Rocky Geyer, and eventually, our research germinated into a large program, the Under Sea Remote Sensing (USRS) Program, to acoustically map different estuaries and their impact on sonar performance. Estuaries can be classified as salt wedge, highly stratified, partially stratified, or well mixed according to the vertical salinity gradients. These different types of estuaries have varying impacts on acoustic scattering, propagation, coherence, and sonar performance.

A persistent feature in many of these estuaries are tidal intrusion fronts, which are surface convergences that typically occur at the mouth of an estuary where fresh, less dense river outflow meets the saltier, denser water. Bubbles are generated and entrained at this front, forming a strong acoustic signal.

Although there have been decades of research on acoustic scattering from bubbles generated by breaking waves, the bubble size distributions in estuarine fronts are strongly modified by the dynamics of the fronts. In addition to mapping estuarine acoustic scattering using broadband echo sounders, I was able to measure the three-dimensional structure and time evolution of shear instabilities, which is typically a challenge because shear instabilities are rapidly evolving. This USRS Program has led to many innovations in how broadband sonars are deployed and has outlined the important role that platforms play in the successful implementation of acoustic technologies.

While my research on estuarine acoustic was going full bore, my colleague Tim Stanton was developing midfrequency broadband acoustic techniques for resonance classification of fish swim bladders for fisheries acoustics applications. Eventually, Tim and I teamed up to continue this development into the deep realms of the ocean.

The ocean twilight zone (OTZ; see twilightzone.whoi.edu) is the vast, globe-spanning layer of water between 200 and 1,000 m depth, home to diverse communities of mesopelagic fishes, cephalopods, crustaceans, and gelatinous organisms. Yet, little is known about the biology, abundance, biomass, distribution, or behavior of these organisms. The OTZ is acoustically characterized by the presence of deep sound-scattering layers (DSL) in shipboard sonar, and, like much of the ocean, it is underexplored and difficult to sample due to a combination of lack of technologies focused on this region, organism patchiness and avoidance, and difficulties capturing fragile species.

Recent evidence suggests that the global OTZ fish biomass may be sufficient to commercially harvest and that much of this biomass performs daily vertical migration (DVM) and may play a critical role in regulating the Earth's climate through the export of carbon to the deep ocean. However, the relative importance of mesopelagic fishes versus zooplankton biomass is still highly uncertain.

Tim and I began the development of an advanced sensor platform, Deep-See (see twilightzone.whoi.edu/deep-see), to fill the technological void for characterizing the OTZ. This towed vehicle integrates wideband, split-beam acoustics (1-500 kHz)

with optical, environmental, and eDNA sensors that can address many of the challenges associated with sampling in the OTZ. We have found that a surprisingly high abundance of organisms can be found outside the dense sound-scattering layers, typically below the noise floor at large ranges for shipboard sonar, and that the target strength of many organisms that perform DVM changes with the depth, which is critical to estimate the biomass.

The WHOI OTZ project is still ongoing and represents one of a number of growing international projects focused on understanding the OTZ, with acoustics playing a central role in exploring and mapping the OTZ. Some of the biggest unanswered questions that can be addressed using acoustic techniques revert back to assessing the importance of the spatial and temporal variability of biomass in the DSL, assessing how much of this biomass participates in daily vertical migration and how much DVM varies in space and time, and to understanding the mesoscale physical drivers of this variability. From a human impact perspective, we need to address questions such as: does the OTZ represent sustainable fisheries, how much does the daily vertical migration of mesopelagic organisms contribute to the global carbon pump, and would mesopelagic fisheries activities adversely impact the ocean's ability to sequester carbon, thus impacting climate.

What are some of the other areas in which you feel you made substantive contributions over your career?

I have been fortunate enough to dedicate the lion's share of my career to research and exploration. However, I consider myself extremely privileged to have also been able to mentor many outstanding graduate students and postdocs. They have enriched my research and expanded my understanding of acoustics and oceanography. Time and again they have taught me that there is still so much left to learn. In thinking about my contributions to acoustical oceanography, I hope time will prove that a substantive contribution lies in the teaching and mentoring of the next generation of acousticians.

What do you think are the most pressing open questions that you would like to focus upon over the next 5-10 years?

We still have a long way to go to fully understand the many scales of variability in the ocean, and acoustic techniques will be center stage in the discovery phase of that journey. Anthropogenic forcing will continue to impact our oceans, with corresponding changes in the propagation, scattering, attenuation, and coherence of acoustic signals, particularly in acoustic hot spots such as in regions of abrupt topography. I am optimistic that combining traditional shipboard platforms, ocean observatories, autonomous underwater and surface vehicles, more compact and less costly acoustic systems, continued theoretical development of physics-based acoustic-scattering models, advanced signal-processing approaches, and harnessing the power of artificial intelligence (e.g., through approaches such as developing machine-learning frameworks for the applications of acoustic classification) will allow us to use acoustics to understand the many scales of temporal and spatial variability in the ocean and to continue to contribute toward ocean conservation and sustainability challenges.

Bibliography

Geyer, W. R., Lavery, A. C., Scully, M. E., and Trowbridge, J. H. (2010). Mixing by shear instability at high Reynolds number. *Geophysical Research Letters* 37(22).

Lavery, A. C., Chu, D., and Moum, J. N. (2010). Measurements of acoustic scattering from zooplankton and oceanic microstructure using a broadband echosounder. *ICES Journal of Marine Science* 67, 379-394.

Lavery, A. C., Geyer, W. R. and Scully, M. E. (2013). Broadband acoustic quantification of stratified turbulence. *The Journal of the Acoustical Society of America* 134, 40-54.

Lavery, A. C., and Ross, T. (2007). Acoustic scattering from double-diffusive microstructure. *The Journal of the Acoustical Society of America* 122, 1449-1462.

Lavery, A. C., Schmitt, R. W., and Stanton, T. K. (2003). High-frequency acoustic scattering from turbulent oceanic microstructure: The importance of density fluctuations. *The Journal of the Acoustical Society of America* 114, 2685-2697.

Lavery, A. C., Wiebe, P. H., Stanton, T. K., Lawson, G. L., Benfield, M. C., and Copley, N. (2007). Determining dominant scatterers of sound in mixed zooplankton populations. *The Journal of the Acoustical Society of America* 122, 3304-3326.

Contact Information

Andone Lavery alavery@whoi.edu

Department of Applied Ocean Physics and Engineering
Woods Hole Oceanographic Institution
Bigelow 211
98 Water Street
Woods Hole, Massachusetts 02543, USA

Sounds Like Research: Graduate Student Stories

Megan S. Anderson, Zane Rusk, Colby Cushing, Lucy Ruoqian Cheng, Hilary Kates Varghese, Mark Langhirt, and Elizabeth Weidner

Acoustic behavior of metamaterials. Oceanic stratification characterization. Foraging behavior of whales. These are just a few of the many research areas explored by our fellow and former members of the Acoustical Society of America (ASA) Student Council (see bit.ly/3GF5jPr). In this essay, we highlight several of our peers who inspire us, all of whom are either nearing graduation or recently graduated. See other *Acoustics Today* essays about students and the Student Council in "*AT* Collections" at bit.ly/3m5kMyP.

Colby Cushing

(see bit.ly/3NOjvYz) graduated in August 2021 with a PhD in mechanical engineering from the University of Texas (UT) at Austin where he is now working as a postdoctoral fellow in the Applied Research Laboratories. He first learned about acoustics as a possible career path from a counselor at Penn State University (University Park, Pennsylvania) who identified acoustics as a connector between Colby's undergraduate major in aerospace engineering and his passion for music. A conversation with Preston S. Wilson at an Acoustical Society of America (ASA) meeting resulted in Colby becoming a graduate student at UT Austin, where he found that his relationships with others in the acoustics program were crucial to his graduate experience. He explains his research as motivated by acoustic metamaterials.

"My PhD research was focused on the characterization of underwater acoustic metamaterials (AMMs). One method of predicting the necessary material properties needed in exotic devices such as underwater acoustic cloaks is spatial transformations (transformation acoustics). These properties can be impossible to realize with standard materials. A metamaterial, with the Greek prefix *meta-* meaning beyond, is one whose dynamic performance goes beyond that of existing materials. The effective properties of AMMs exceed what is typically possible with standard mixing laws of composite materials by leveraging both subwavelength structures and composition and can be used to physically realize the material properties predicted by transformation acoustics. However, there are often complex geometries that can require state-of-the-art fabrication techniques (e.g., additively manufactured metal structures) and are especially difficult to experimentally characterize for underwater applications. In my work, I used new and existing experimental apparatuses to characterize specific underwater AMMs. I also explored the relationships between homogenization, numerical simulation, and fabrication methods."

Colby's current research uses the fundamental ideas in his work with metamaterials to examine the substructures present within seagrass leaves and their acoustic effects in underwater environments.

Lucy Ruoqian Cheng

(see lucyrc.github.io) will be graduating with a PhD in linguistics from the University of Kansas (Lawrence) this year. Lucy's undergraduate degree in the Chinese language from Fudan University (Shanghai, China) was only the start of her scholarship in linguistics, with her research now spanning colloquial Beijing, Mandarin, English, and Eastern Min (Fuzhou) languages. Lucy was drawn to acoustics while studying ongoing sound change across speakers, and she explores the effect of age on speech perception in her dissertation.

"The way listeners process speech sounds changes across the life span, even when the hearing acuity has not yet started to decline. A better understanding of the aging effect in its early stage (55-65 years old) contributes to the advancement of speech technology (e.g., hearing aids and voice/user interfaces) to better fulfill the needs of an aging population. My doctoral dissertation investigated the influence of aging on speech perception by comparing speech perception

and auditory perceptual plasticity in older versus younger adults with normal hearing. The experiments investigated the listeners' use of the acoustic signal and contextual information and tested how listeners adapt to ambiguity in speech sounds. Younger listeners exhibited greater perceptual plasticity than older listeners when adjusting their acoustic cue weighting, whereas both younger and older listeners demonstrated good use of contextual information."

In addition to her research, Lucy's care for people and language has led her to volunteer on a project developing Audio Speech Recognition training data and models for Eastern Min, her native tongue.

Hilary Kates Varghese graduated with a PhD in oceanography from the University of New Hampshire (Durham) in December 2021. Her journey into bioacoustics began on an undergraduate field trip where she learned that whales sang in the New York, NY, harbor and that these whales were studied by scientists. She knew then that she wanted to be one of those scientists! After completing her degree in biological sciences at Cornell University (Ithaca, New York) and taking advantage of undergraduate research opportunities along the way, Hilary gained real-world work experience at an algal biofuel company while getting a master's degree in applied mathematics at Florida Gulf Coast University (Fort Myers). This provided her with the critical thinking and quantitative analysis skills needed to pursue her interests in understanding how human-generated sound affects marine life.

"My PhD research looked at the potential interaction between 12-kHz deepwater multibeam ocean-mapping activity and Cuvier's beaked whale foraging behavior. To the best of current understanding, this scenario represents a potential worst case for the interaction of ocean-mapping sonar with marine life due to the overlap in the use of this sound in areas where these animals live as well as in the frequency overlap of the sound with beaked whale hearing. To assess the potential effect, I performed spatial and temporal analyses on the echolocation clicks of beaked whales, used as a proxy of their foraging behavior, collected from an array of hydrophones spanning an 1,800-square-kilometer area. Changes in foraging behavior were assessed before, during, and after two 12-kHz multibeam mapping surveys conducted over the array. The results of my research showed that beaked whales in the area did not stop foraging and did not leave the area during the two surveys, revealing that the mapping activity did not adversely affect beaked whale foraging."

Hilary always hoped that her research would influence marine policy and noise regulation in the ocean, and she strengthened her science communication skills during an internship with *Acoustics Today* (see bit.ly/3Q0OVgy). Now, in her role as a marine bioacoustician with the Bureau of Ocean Energy Management (a part of the US Department of the Interior), she is putting her expertise to use everyday.

Mark Langhirt (see bit.ly/3t8txMw) graduated with a PhD from Penn State University (State College, Pennsylvania) this year. While working as a civilian on autonomous underwater robots with naval sonars, he became convinced that "acoustics is one of the best tools we have for studying the ecosystems and dynamics that shape the hidden world beneath the waves." This experience, combined with his undergraduate degree in physics, motivated Mark to gain a deeper theoretical understanding of underwater environments by studying acoustics at the graduate level. Mark focused his research on the following question: How can we construct an underwater acoustic propagation model that is less computationally expensive than the alternatives and capable of capturing primary effects we see in three-dimensionally inhomogeneous underwater environments?

"My research focuses on deriving and implementing 3D underwater acoustic propagation models based on energy conservation principles. The energy flux method integrates modal acoustic energy directly without solving for eigenvalues, and therefore takes significantly less time to execute compared to other propagation models. My contribution has been to extend the 2D energy flux theory and derive semi-coherent 3D energy flux models that capture horizontal refraction and focusing in underwater environments like shorelines and canyons. The models are verified by comparison with both analytical solutions and other

3D acoustic propagation models. As far as we are aware, these are the first semi-coherent 3D energy flux models that have been developed. These models could be beneficial in autonomous remote sensing networks with limited computational resources or in scenarios with uncertain acoustic environmental parameters."

After graduation, Mark hopes to continue working in ocean sciences research, with either an independent or government-affiliated research lab.

Elizabeth Weidner

(see bit.ly/3M5lXJb) graduated this summer with a joint PhD from the University of New Hampshire (Durham) and Stockholm University (Stockholm, Sweden). She first became "hooked" on acoustical oceanography when, as an undergraduate, she learned that "sound could provide observations of the deep ocean where light could not." As a geophysicist in the private sector, she used acoustic systems to map the water column and seafloor, developing an increasing interest in the impact of climate change on the Arctic Ocean. These experiences ultimately led Elizabeth to enroll as a graduate student at the Center for Coastal and Ocean Mapping (CCOM) at the University of New Hampshire where she asks: How can we leverage broadband acoustic data to better understand processes within the ocean water column (and seafloor)?

"My PhD research is focused on the acoustic analysis of oceanic density stratification structure utilizing broadband split-beam echo sounders. Stratification structure is a ubiquitous feature in the world's oceans and influences the vertical transport of heat and many dissolved constituents (e.g., nutrients, carbon, oxygen) in the water column. While scientists have been using acoustic systems to observe scattering from ocean structure for many decades, my work has leveraged the high SNR, vertical range resolution, and frequency-modulated scattering of broadband systems to move beyond just observation to tracking, analysis, and acoustic inversion for the estimations of physical water column properties. I hope my work will be used to better understand evolving spatiotemporal changes in ocean stratification from climate change."

After defending her dissertation in both New Hampshire and Stockholm, Elizabeth will work as a research scientist at the CCOM before starting a postdoc in early 2023.

Overview

Although these graduate students pursued very different research topics, they all credit participation in the ASA as contributing to their professional and personal success. Elizabeth identified service to the community as a "hugely important part of career development," an idea that Hilary echoed. They also discussed, along with Lucy and Mark, the value in attending ASA meetings, presenting your work, and meeting other researchers. Colby also reflected on how inspiring ASA meetings are, even as he noted the imposter syndrome that can sometimes come when you're surrounded by impressive researchers. He joined the ASA Student Council hoping "to help new students continue to feel welcomed in the Society but not overwhelmed by the gravity of the organization so that they can feel comfortable making an immediate positive impact." The impact of these individuals has certainly been felt on the ASA Student Council, and we look forward to seeing where their careers take them.

Contact Information

Megan S. Anderson andersonm@gwu.edu

The George Washington University
800 22nd Street NW
Suite 3000
Washington, DC 20052, USA

Zane Rusk ztr4@psu.edu

Department of Architectural Engineering
The Pennsylvania State University
104 Engineering Unit A
University Park, Pennsylvania 16802, USA

Colby Cushing colby.cushing@utexas.edu

Lucy Ruoqian Cheng rqcheng@ku.edu

Hilary Kates Varghese hsk33@cornell.edu

Mark Langhirt mark.a.langhirt@gmail.com

Elizabeth Weidner eweidner@ccom.unh.edu

High-School Students Win ASA Awards at the International Science and Engineering Fair

Laurie M. Heller

Annually, the Acoustical Society of America (ASA) is one of the sponsors of the Regeneron International Science and Engineering Fair (see bit.ly/3ykRalG). This year's meeting took place in Atlanta, Georgia from May 8-13, 2022, and included 1,750 projects. The ASA reviewed all the acoustics projects and is delighted to share with members the work of the four exceptional high-school students who were awarded prizes by the Society.

High-school student **Amara Orth** (Lewis Central High School, Council Bluffs, Iowa) used her family's honeybee farm as her pandemic science lab. Her scientific curiosity was piqued when her grandfather claimed that he could discern the health of a beehive by its sound. This culminated in her award-winning science fair project that showed that vibroacoustics are indicators of bee health and that hidden Markov models can accurately characterize the important health states of honeybee colonies. New to both acoustics and machine learning at the start of her project, Amara began her research by searching for appropriate piezoelectric microphones to record the beehives' vibroacoustic signals. She recorded dozens of hives over a three-month period and utilized her beekeeping expertise to classify each hive into one of nine possible health states (e.g., the loss of a queen or volatile chemical exposure). After examining spectral features of the recordings, she trained a hidden Markov model to classify new waveforms into the correct health states with 92% accuracy. Amara's discovery could help beekeepers use remote monitoring and intervene early to prevent colony collapse. The ASA awarded her a First Prize of $1,500 (plus $200 for her school and $500 for her mentor, Mrs. Shannon).

The ASA awarded Second Prize to **Chinmayi Ramasubramanian** (Sri Kumaran Children's Home, Bangalore, India) for her application of acoustic analysis to elephant vocalizations to create an early detection system for elephants near human populations to reduce human-elephant conflict. Her unique signal processing was used in a machine-learning model that could classify recordings of elephant vocalizations as a chirp, roar, rumble, or trumpet. She installed a small alarm system containing her neural network model implemented in real time on a Raspberry Pi microcomputer along with a unidirectional microphone and a system that sends a telegram when nearby elephants are identified. The Second Prize was $1,000 (plus $100 for her school and $250 for her mentor).

The Third Prize was awarded by the ASA to **Gwyneth Liu** (Mills E. Godwin High School, Henrico, Virginia) for finding ways to improve the efficiency of a wave energy converter via her model of adaptive control. Her adaptive model consisted of a spectral analysis-based sea-state classification algorithm that found that the swash plate angle ratio was an important parameter. The goal is to convert the kinetic energy of hydraulic wave movement into usable electric energy. The Third Prize was $600 (plus $150 for her mentor).

Honorable Mention went to **Anu Iyer** (Little Rock Central High School, Little Rock, Arkansas) for using the acoustics of brief vocalizations to classify patients with Parkinson's disease. Her approach

used very short utterances that were recorded on voice mail.

In addition to cash prizes, all awardees were invited to attend an ASA meeting, with a waived registration fee (for the student plus a mentor) and a partial contribution toward travel costs. The awardees' abstracts can be viewed at exploresound.org/isef-asa-winners.

The ASA judging team was led by Laurie Heller (Carnegie Mellon University, Pittsburgh, Pennsylvania) and included Jeffrey Vipperman (University of Pittsburgh, Pittsburgh, Pennsylvania), Robert Smith (Penn State University, University Park, Pennsylvania), and Andy Chen (Matters Academy, London, United Kingdom). The judges commend the accomplishments of the many talented youth who participated in the Regeneron International Science and Engineering Fair!

Contact Information

Laurie M. Heller laurieheller@cmu.edu

Carnegie Mellon University
Pittsburgh, Pennsylvania 15217, USA

The Journal of the Acoustical Society of America

JASA Call For Submissions:

JASA is currently accepting manuscripts for the following Special Issues:

- Perception and Production of Sounds in the High-Frequency Range of Human Speech
- Fish Bioacoustics: Hearing and Sound Communication

Special Issue articles are free to read for one year after publication and don't incur any mandatory page charges.

Find out more at
asa.scitation.org/jas/info/specialissues

POMA Student Paper Competition

ATTENTION STUDENTS ATTENDING THE NASHVILLE MEETING!!

Proceedings of Meetings on Acoustics (*POMA*) will once agin be holding a society-wide student paper competition for a *POMA* submission based on a presentation or poster from the ASA meeting in Nashville this spring.

Award Amounts: Up to five student papers will receive an award of USD $300.

For qualification requirements, submission window, and additional information, please visit:
acousticalsociety.org/asa-meeting
or contact the POMA Office at
poma@acousticalsociety.org

FOLLOW THE ASA ON SOCIAL MEDIA!

@acousticsorg

@acousticsorg

The Acoustical Society of America

AcousticalSociety

AcousticalSocietyofAmerica

acousticstoday.org/youtube

Online Education Tools: Effective Practices from DOSITS

Kathleen J. Vigness-Raposa, Holly Morin, Christopher Knowlton, and Gail Scowcroft

Online learning through virtual seminars (webinars), live streaming, virtual meetings, website material, and instructional videos all rose in use and popularity during the 2010s. With the onset of the Covid-19 pandemic in 2020, the Acoustical Society of America (ASA) and its members, along with the rest of the world, dramatically increased their use of these virtual tools. Life was transformed into Zoom meetings, Microsoft Teams engagements, and other online learning experiences. This cultural shift has led to some positive developments, such as broader engagement in learning experiences with those whose involvement has been limited previously by a lack of travel funding. However, there have been some negatives, including calendars overcommitted with many more online learning experiences that lacked the depth one would experience in person.

Still, even with some negatives, the very positive advantages of online learning and meetings suggest that even as the pandemic wanes, these new-found tools and skills will not. Indeed, the ASA, like many other societies, has started to explore how it can expand its impact by providing educational tools that not only are directed at members but also at a wider community interested in acoustics in all of the diverse ways in which it is approached by its members.

Moreover, along with the increased use of virtual tools, the selections of available tools have increased, and the tools themselves have been greatly improved. Most virtual platforms are now more stable, are easier to use, and have added functionality, and perhaps most significantly, most people have become more familiar with using such tools, particularly as participants in events or consumers of information.

Knowing about the increased interest in online education and tools by the ASA and its members, the editor of *Acoustics Today*[1] asked the Discovery of Sound in the Sea (DOSITS) project, an affiliate of the ASA and *Acoustics Today*, to share lessons learned over the 20 years of our online educational presence.

As background, Scowcroft (2021) provides the history of the DOSITS project, which celebrated its 20th anniversary last year and will be celebrating the 20th anniversary of the project website (see dosits.org) on November 30, 2022, with a special webinar (see dosits.org/webinar2022). One of our most cherished honors is the Science Writing Award in Acoustics for Media other than Articles presented by the ASA in 2007. The foundational core of the project is to make acoustics accessible for all audiences, including the public, educators, students, news and media professionals, and decision makers (Vigness-Raposa et al., 2014, 2016). DOSITS has accomplished this using multiple digital platforms (Morin et al., 2017) and educational principles that develop critical thinking skills and deeper engagement.

Thus, the purpose of this article is to share best practices for developing online learning experiences, highlighting the techniques for building and engaging your audience. It provides insight into some educational tools that have been very successful in engaging and challenging participants in their online learning. Multiple digital platforms as potential vehicles for online learning are presented, including webinars and how to structure them to have broad reception and impact; websites with sufficient versatility to engage across different devices and guide participants along a progressive path of knowledge development; and short instructional videos to build interest and animate abstract or hard-to-visualize scientific processes.

---

1 For full disclosure, the editor of *Acoustics Today* is a member of the DOSITS science advisory board.

Building and Engaging Your Audience

Online learning experiences should be designed to target a specific audience and meet its needs while also understanding the obstacles and challenges for online learning. Needs assessments can be helpful, asking your audience members about their backgrounds, learning expectations, content needs, and engagement characteristics. Understanding their foundational knowledge and the content and format they find most useful to support their work is critical to building and engaging your audience.

The content of the experience should also be targeted. Are you trying to fill a knowledge gap in existing resources, providing expanded information on a specific topic or are you trying to present new information to an informed audience? Because online audiences may come from a variety of backgrounds, it is important to provide links to resources covering foundational information, designing the learning experience to bring all audience members to the same level, allowing for successful engagements.

Part of knowing your audience is a consideration of logistics. If you are expecting to reach a global audience in real time, you must carefully select a date and time that is most convenient to the majority of participants. Are there technological challenges, such as times at which Internet connectivity or accessibility of certain online tools are reduced? It is important to remember that participants may not have administrative access to their computers and installing novel software could be difficult.

Finally, you must market your online learning experience to your potential audience. Make sure there are posts on common email listservs, appropriate professional society bulletins, or perhaps social media platforms. Providing information about the content and format of the event is critical. It may be helpful to develop an outline or create a short video to advertise the experience. Ensuring that the audience's expectations are aligned with the learning experience is foundational to success.

Specific Tools
Webinars and Other Live Events

This discussion focuses on webinars, but other live events, such as live streaming, can be related depending on how they are structured. Certain audiences are unable to join Webex or Zoom webinars because of security restrictions.

Thus, it is advantageous to provide multiple options for the audience to participate in the webinar, such as concurrent live streaming on YouTube or Facebook.

The first step in designing a successful webinar is to select an appropriate speaker or presenter. This individual should be knowledgeable in the focused content of the webinar as well as in the goals of the webinar and background information on the audience. The speaker should also conduct a trial run of their presentation to ensure compatibility of their media with the software platform as well as with the appropriate Internet bandwidth, lighting, and quality of their audio. During the trial run, feedback to the speaker is helpful on topics such as pace of delivery, clarity of terms, and progressive development of knowledge and/or references to background information.

The webinar presentation should be structured to include references to ongoing research and/or case studies as well as an interactive question and answer (Q&A) session. Audiences appreciate the real-world connection provided by case studies that demonstrate the implementation of the science being presented. For the Q&A session, the audience should submit questions to the webinar platform throughout the presentation. A moderator can curate questions and pose them to the presenter, providing for real-time interactions with the expert. We have found that webinars should be about 1 hour in duration, with a presentation of approximately 40 minutes and a Q&A session of 20 minutes. This allows enough time for the speaker to present their topic in sufficient depth without being tempted to dive too deep, for the audience to ask questions and engage with the presenter, and for participants to find time for the webinar in their schedules. During the Q&A session, questions should be included from all participation platforms.

The webinar should be supported with associated resources to engage the audience. Providing an outline of the webinar to participants prior to the event is highly recommended. The outline should include Internet links to background information and/or scientific papers to enhance and elaborate on the webinar concepts. After the webinar, a PDF of the presentation and a transcript of the event can provide deeper audience engagement. Recording the webinar to allow for asynchronous viewing is encouraged, but this may also require some technical editing prior to posting.

Figure 1. *The web page for a Discovery of Sound in the Sea (DOSITS) webinar demonstrating best practices for audience engagement (see dosits.org/modeling).*

Figure 1 is a screen grab illustrating these principles for a recent webinar hosted by the DOSITS project. The webinar was "Acoustic Propagation Modeling" presented by Aaron Thode of the Scripps Institution of Oceanography (La Jolla, California) (see dosits.org/modeling). An outline of the webinar was provided prior to the event with links to supporting content on the DOSITS website. After the webinar was conducted, a PDF of the presentation and a transcript of the webinar were created and posted. In addition, an archived recording of the webinar was made available. After the webinar is conducted, a survey of the participants can provide very constructive, immediate feedback on the level of the content, the pace of the presentation, and whether the needs of the audience were met as well as providing suggestions for future webinar topics and/or speakers. It may also be important to your audience to be able to receive a Certificate of Professional Development or verification of participation. The DOSITS project has conducted numerous needs assessments of our audiences, and this has been continually mentioned as a much-appreciated benefit to our annual webinar series.

Websites

Websites were one of the first available online learning tools, but their structure and function have changed considerably over the past two decades as Internet connectivity has increased. However, the foundational principles of good online learning websites remain steadfast. The content should be organized into digestible sections, with titles and structure that challenge inquiry. Starting titles with "what," "how," or "why" challenges the audience to think critically and internalize the content. Making the content interactive also leverages the advantages of online tools. For the DOSITS project, the Audio Gallery (see dosits.org/audio) and the Science of Sound (dosits.org/science) are favorite sections because their interactive nature targets a wide variety of interests and ability levels.

Science topics may have a certain level of controversy, so basing a website on peer-reviewed literature is critical. The content of the website is verified and validated by providing references and links to peer-reviewed material. Synthesizing published science into digestible content for an audience can be a focused project goal, such as it is for DOSITS. Using peer-reviewed literature ensures that the scientific community has approved of the science and its methods. To further ensure that the synthesis process has not compromised the scientific integrity of the peer-reviewed literature, peer review of the website content itself can be conducted. The DOSITS project holds semiannual Advisory Panel meetings to review all content before it is published. Advisory Panel members, all of whom are ASA members, also work together to ensure that relevant publications are included in DOSITS content. It is also important to be transparent about how the website is funded and who has built it and is responsible for its maintenance.

Finally, a best practice for educational websites is to create structured tutorials. Structured tutorials link successive web pages to progressively develop knowledge on a particular topic. The web structure of websites can be overwhelming for some audience members, so a structured tutorial provides a clear starting point and a path for incrementally developing knowledge. For example, the DOSITS project has developed three structured tutorials: The Science of Sound, The Effects of Underwater Sound on Marine Life, and The Technologies of Underwater Sound (see dosits.org/tutorials).

Instructional Videos

Instructional videos should be short, focused clips that tap into the YouTube generational mindset of

instructional learning. It is important that instructional videos are not just lectures or talking heads but that they include animations, illustrations, or field recordings, leveraging the advantages of online learning tools. Video length will depend on the content being presented, but a general rule is a duration of less than five minutes. If more time is needed to adequately address the content, then a series of instructional videos would better serve the audience. Also make sure the instructional video is staying focused on a particular topic. Again, if the topic is complex, a series of instructional videos, each one focused on one facet of the content, would engage the audience better. The DOSITS project has developed three instructional videos: Science of Sound, Marine Mammal Hearing, and Hearing in Marine Fishes (see dosits.org/videos).

Summary

Online learning is a great mechanism for reaching diverse audiences, and its implementation will continue in popularity. Simple tweaks to existing practices can help build and engage audiences, propelling your online teaching to greater rewards. Remember that online learning is often optional; the audience is not required to complete a class or project as is often the case with traditional learning. Because online learning is something audience members are choosing to do, conducting questionnaires or needs assessments will inform your content choices and selection of appropriate educational tool(s). By understanding the interests and preferences of your audiences, you will develop educational tools that attract audiences as well as educate them.

References

Morin, H., Vigness-Raposa, K. J., Knowlton, C., Scowcroft, G., Miller, J. H., Ketten, D. R., and Popper, A. N. (2017). Implementing multiple digital platforms to effectively communicate research on underwater acoustics. *Proceedings of Meetings on Acoustics* 27(1), 010007. https://doi.org/10.1121/2.0000255.

Scowcroft, G. (2021). The Discovery of Sound in the Sea project: Twenty years of success in synthesizing science for nonexperts. *Acoustics Today* 17(4), 78-80. https://doi.org/10.1121/AT.2021.17.4.78.

Vigness-Raposa, K. J., Scowcroft, G., Miller, J. H., Ketten, D. R., and Popper, A. N. (2016). Discovery of Sound in the Sea: Resources for educators, students, the public, and policymakers. In Popper, A. N., and Hawkins, A. (Eds.), *The Effects of Noise on Aquatic Life II*. Springer-Verlag, New York, NY, pp. 1183-1190.

Vigness-Raposa, K. J., Scowcroft, G., Morin, H., and Knowlton, C. (2014). Underwater acoustics for everyone. *Acoustics Today* 10(2), 50-59.

Contact Information

Kathleen J. Vigness-Raposa
kathy@INSPIREenvironmental.com

INSPIRE Environmental
513 Broadway, Suite 314
Newport, Rhode Island 02840, USA

Holly Morin holly_morin@uri.edu

Christopher Knowlton cknowlton@uri.edu

Gail Scowcroft gailscow@uri.edu

Graduate School of Oceanography
University of Rhode Island
Narragansett, Rhode Island 02882, USA

The Journal of the Acoustical Society of America
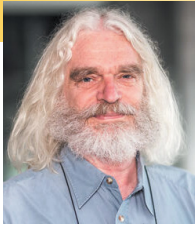
Reflections

Don't miss Reflections, *The Journal of the Acoustical Society of America's* series that takes a look back on historical articles that have had a significant impact on the science and practice of acoustics.

See these articles at:
acousticstoday.org/forums-reflections

Obituary
Armin Kohlrausch, 1953–2022

....

Armin Kohlrausch, a Fellow of the Acoustical Society of America, passed away unexpectedly on March 25, 2022, shortly after the meeting of the German Acoustical Society in Stuttgart (Germany). Armin studied acoustics in Goettingen, Germany, and received his PhD under the supervision of Manfred Schroeder on topics related to binaural hearing (Kohlrausch, 1985). In 1991, he joined the Institute for Perception Research at the Eindhoven University of Technology (UT/e; Eindhoven, The Netherlands) to continue his research on hearing, with a focus on modeling the effective signal processing of the auditory system. During that time, many of his influential publications on modeling of auditory modulation perception and binaural perception emerged (Dau et al., 1997). One key aspect of the modeling work was to model the limits of auditory perception by making a detailed description of the early stages of auditory signal processing that would represent the information reduction along the auditory pathway, followed by an optimal detector framework representing more central processing, an approach that allowed modeling a wide range of perceptual phenomena related to auditory masking and discrimination performance.

In 1999, Armin moved to the Philips Research Laboratories in Eindhoven while maintaining a part-time affiliation with the TU/e. Here he built up a new research group focusing much more on applied topics such as low-bitrate audio coding. The basic knowledge built up at the Institute for Perception Research turned out to be a very fruitful basis for contributing new methods to parametrically encode spatial parameters of audio signals in this way, allowing for much more efficient low-bitrate audio-encoding algorithms (Breebaart et al., 2005; Kohlrausch, 2007). This eventually led to contributions to the Moving Picture Experts Group (MPEG) surround standardization. In addition, his research focused on various applied topics such as music information retrieval and, later, soundscapes in intensive care units. Armin maintained a strong interest in basic research, and as part of his affiliation with the TU/e, he contributed to our understanding of the factors that influence auditory-visual synchrony perception.

Armin's scientific curiosity extended far. Among his papers are contributions to the perception of room acoustics, the irrelevant speech effect, and the analysis of soundscapes in an intensive care unit and on music information retrieval and the history of psychoacoustics. One of his recent endeavors was the rediscovery of the work of Alvar Wilska from 1938 that was a pioneering work on spatial perception and included listening experiments that used a dummy head with accurate anthropomorphic features. Armin was awarded the Helmholtz Prize from the German Acoustical Association (DEGA) in 2017.

During his career, Armin contributed to many technical committees and was a devoted mentor to many young scientists in our field. As a PhD supervisor, he had a wonderful talent to give considerable space to his students in developing their ideas, although he also probed their work with his unsurpassed analytical skills to improve the scientific strength of their work. At the same time, Armin was a truly supportive and kind mentor, helping young inexperienced students to become confident and well-recognized scientists.

Selected Publications by Armin Kohlrausch

Breebaart, J., Kohlrausch, A., van de Par, S., and Schuijers, E. (2005). Parametric coding of stereo audio. *EURASIP Journal on Applied Signal Processing* 9, 1305-1322.

Dau, T., Kohlrausch, A., and Kollmeier, B. (1997). Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers. *The Journal of the Acoustical Society of America* 102(5), 2892-2905.

Kohlrausch, A. (1985). Psychoacoustic studies of spectral aspects of binaural hearing. *The Journal of the Acoustical Society of America* 74, 1441.

Kohlrausch, A. (2007). The perceptual basis for audio compression. *Physics Today* 60(6), 80-81.

Written by:
Aki Härmä aki.harma@philips.com
   Philips Research, Eindhoven, The Netherlands

Steven van de Par steven.van.de.par@uol.de
   Carl-von-Ossietzky Universität, Oldenburg, Germany

Obituary
Sam H. Ridgway, 1936–2022
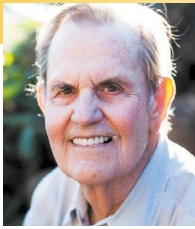
Sam H. Ridgway, pioneering marine mammal veterinarian and bioacoustician, passed away on July 10, 2022, in San Diego, California. Sam was born in Bigfoot, Texas, in 1936. He earned his undergraduate degree in 1958 and a veterinary degree in 1960 from Texas A&M University (College Station). Sam, a commissioned veterinary officer in the US Air Force, was transferred to Point Mugu, California, in 1960 where he became one of the founders of the US Navy Marine Mammal Program (MMP). In 1970, he received a Navy fellowship to study under Richard Harrison at Cambridge University (Cambridge, United Kingdom) where he earned a PhD in neurobiology.

In 2007, Sam became founding president and CEO of the National Marine Mammal Foundation. Among his awards, Sam was elected Fellow of the Acoustical Society of America. He also received both the Lifetime and Clinical Medicine Awards from the International Association for Aquatic Animal Medicine and the Kenneth S. Norris Lifetime Achievement Award from the Society for Marine Mammalogy.

Sam combined his interests in basic and applied science to advancing marine mammal medicine, bioacoustics, physiology, and behavior. He was known for his pioneering and ground-breaking discoveries that advanced the health and welfare of dolphins, porpoises, whales, sea lions, and seals. Indeed, often called the "father of marine mammal medicine," Sam worked tirelessly in developing and refining his science.

Sam began his acoustic studies in the laboratory of E. G. Wever at Princeton University (Princeton, New Jersey). During those studies, Sam and his colleagues did the first recording of evoked potentials to measure turtle hearing as well as a series of studies that provided the first detailed examination of the dolphin ear and peripheral auditory system using anesthetized animals. Sam's work with evoked potentials helped set the groundwork for their broad use in the study of marine mammal hearing today, including pioneering efforts to study the hearing of stranded cetaceans for which no prior hearing information exists.

Sam was a scientist of many firsts, being the first to experimentally study temporary threshold shifts in cetaceans, the first to utilize functional medical imaging to study dolphin hearing, and the first to study the impact of hydrostatic pressure due to diving on the hearing of toothed whales.

Sam published more than 350 papers, books, and book chapters, and he wrote a popular book, *The Dolphin Doctor*. Although prolific in his work and a giant in the field, Sam was the quintessential mentor, teacher, and friend. One of the hallmarks of Sam's career was his devotion to both veterinary and doctoral students. Indeed, to his last days, Sam always found time for students.

Sam always managed his life with kindness and refinement. He was acknowledged worldwide and deeply admired. Sam's wife of almost 60 years, Jeanette Fuller Ridgway, passed away in 2020. Sam is survived by his brothers Don Ridgway and Sid Ridgway and their families.

Selected Publications by Sam H. Ridgway

Ridgway, S. H. (2008). History of veterinary medicine and marine mammals: A personal perspective. *Aquatic Mammals* 34, 471-513.

Ridgway, S. H. (2019). Ask an acoustician: Sam H. Ridgway. *Acoustics Today* 15(2), 45-47. https://doi.org/10.1121/AT.2019.15.2.45.

Ridgway, S. H., Wever, E. G., McCormick, J. G., Palin, J., and Anderson, J. H. (1969). Hearing in the giant sea turtle, *Chelonia mydas*. *Proceedings of the National Academy of Sciences of the United States of America* 64, 884-890.

Wever, E. G., McCormick, J. G., Palin, J., and Ridgway, S. H. (1971). The cochlea of the dolphin, *Tursiops truncatus*: Hair cells and ganglion cells. *Proceedings of the National Academy of Sciences of the United States of America* 68, 2908-2912.

Written by:

Patrick W. Moore patrick.moore@nmmf.org

Cynthia R. Smith cynthia.smith@nmmf.org

Dorian S. Houser dorian.houser@nmmf.org

National Marine Mammal Foundation, San Diego, California

Obituary
Sigfrid D. Soli, 1946–2022

Sigfrid D. (Sig) Soli, an internationally respected scientist in speech perception, cochlear implant research, and hearing aid research, passed away on April 11, 2022, after a short illness. He was 75 years old.

Born in Granite Falls, Minnesota, in 1946, Sig graduated from St. Olaf College (Northfield, Minnesota) in 1968, followed by service in the US Air Force. He obtained his PhD in experimental psychology from the University of Minnesota (Minneapolis) in 1979, followed by a faculty position in psychology at the University of Maryland, College Park (College Park).

In 1984, Sig took a position on a team at the 3M Corporation (Maplewood, Minnesota) to develop a new implantable hearing device, the cochlear implant. He and colleagues published a notable paper (Van Tasell et al., 1987) in which they showed that listeners could recognize 40% of consonants correctly, even when all spectral information was removed. This finding was highly surprising and helped explain the good outcomes of early patients with single-channel cochlear implants.

In 1989, Sig joined the House Ear Institute (Los Angeles, California) to continue his work on cochlear implants and hearing aids. Sig and his team produced the Hearing in Noise Test (HINT) (Nilsson et al., 1990) to evaluate the effects of interfering noise on speech recognition before and after clinical interventions. The HINT is now the international standard for quantifying the effects of noise on speech recognition and has been translated and validated in more than 20 languages.

Sig and his colleagues developed a series of tests to evaluate functional directional hearing (Soli et al., 2018). This test is now used internationally to evaluate localization ability in many professionals (police, fire, and medical) who often must make time-critical decisions based on the direction of sounds indicating, for example, danger or cries for help.

Sig was an advocate for sound-quality standards for classrooms to help children with hearing impairments. He demonstrated that many types of noise present in most classrooms can lead to serious problems in the development of speech and language (Nelson et al., 2002).

Overall, Sig realized the importance of basic science to identify the underlying causes of hearing disorders, working with companies to develop commercial products and with clinicians to apply those products to help patients with hearing disorders: bench to commercialization to patients.

Sig was generous with his time in providing public service to the Acoustical Society of America as a chair of its Speech Communication Technical Committee and was named a Fellow. On a national scale, he served as an advisor to the US National Institute on Deafness and Other Communication Disorders of the National Institutes of Health and to the US National Research Council. He was a voting member of the Ear, Nose, and Throat Device Panel of the US Food and Drug Administration.

In addition to his scientific achievements and public service, Sig was a friendly figure who will be remembered fondly by many colleagues around the world. Sig is survived by his wife Susan and sons Andrew and Daniel.

Selected Publications by Sigfrid D. Soli

Nelson, P., Soli, S., and Seltza, A. (2002). *Acoustical Barriers to Learning: Understanding the Need for a Classroom Acoustics Standards*. Acoustical Society of America, Melville, NY.

Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America* 95, 1085-1099.

Soli, S. D., Giguere, C., Laroche, C., Vaillancourt, V., Dreschler, W. A., Rhebergen, K. S., Harkins, K., Ruckstuhl, M., Ramulu, P., and Meyers, L. S. (2018). Evidence-based occupational hearing screening I: Modeling the effects of real-world noise environments on the likelihood of effective speech communication. *Ear and Hearing* 39, 436-448.

Van Tasell, D. J., Soli, S. D., Kirby, V. M., and Widin, G. P. (1987). Speech waveform envelope cues for consonant recognition. *The Journal of the Acoustical Society of America* 82, 1152-1161.
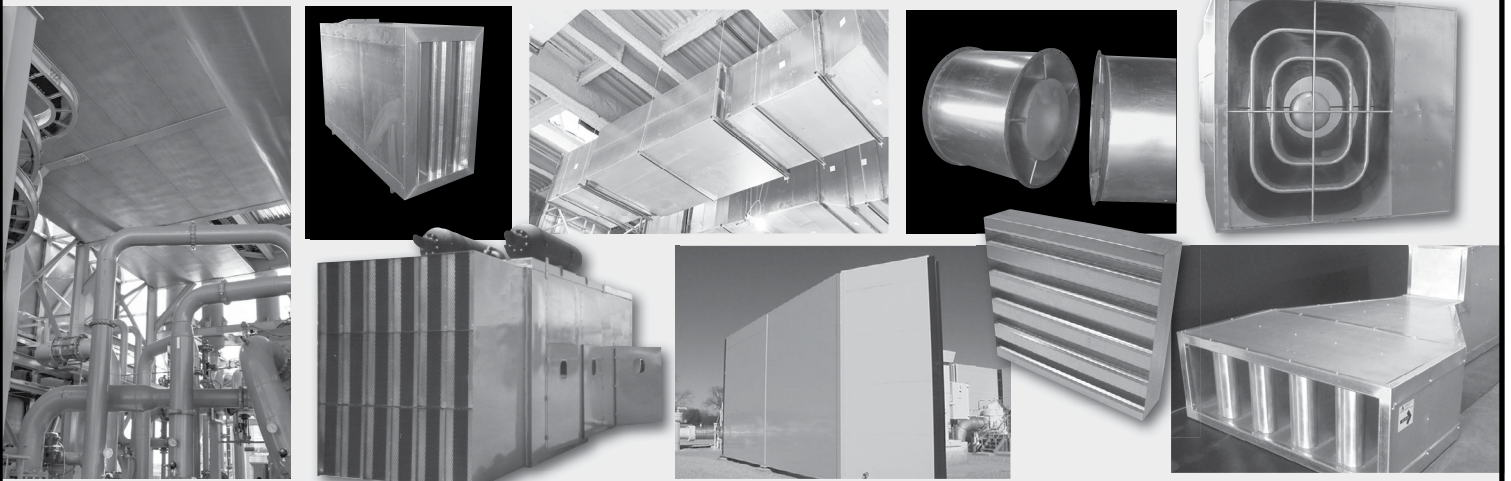
Written by:
Robert Shannon robertvshannon@gmail.com
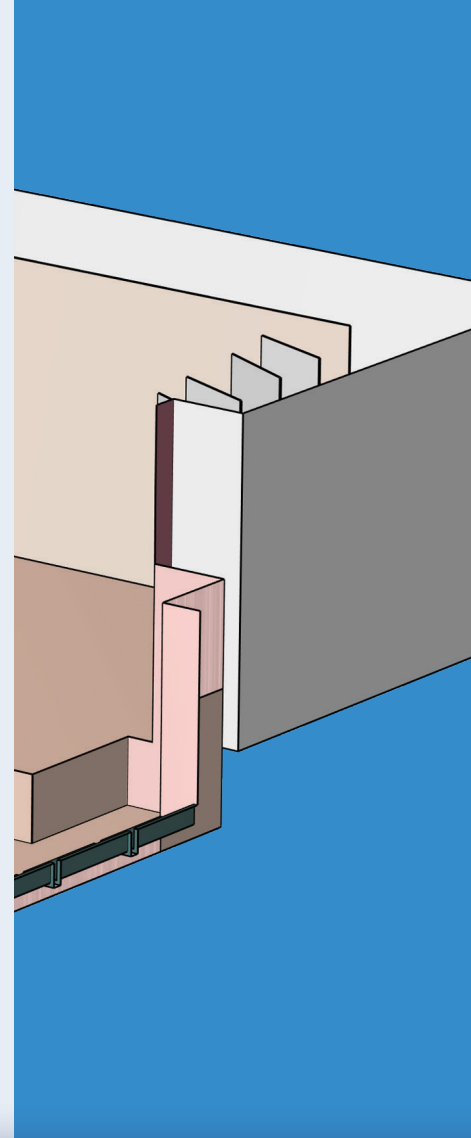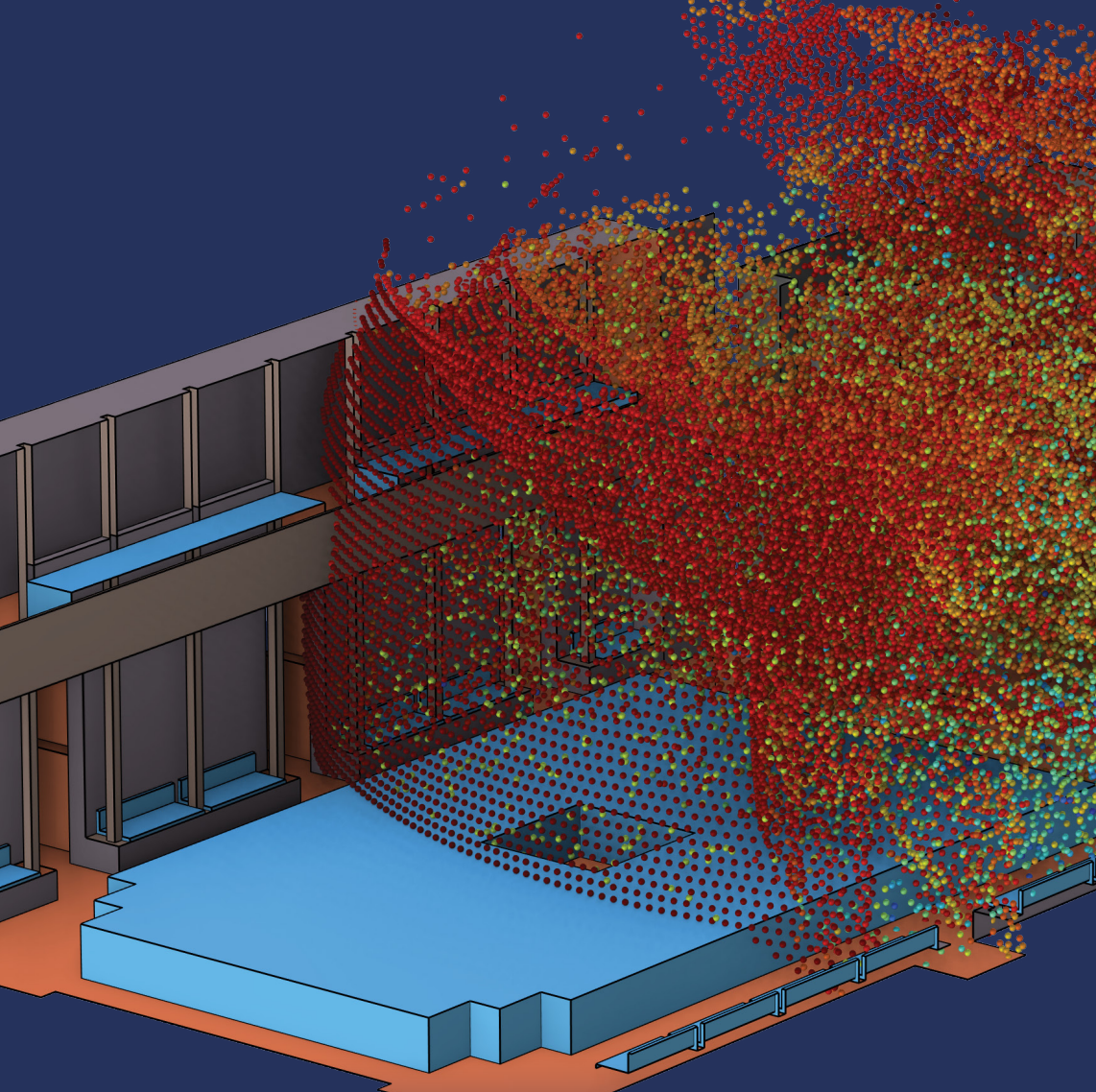
Take the Lead in Acoustics

with COMSOL Multiphysics®

Multiphysics simulation drives acoustics innovation by providing insight into every design aspect that influences product performance. The ability to account for coupled physics phenomena lets you predict, optimize, and virtually test a design under real-world conditions – even before a first prototype is built.

» **comsol.com/feature/acoustics-innovation**

COMSOL