Acoustics Today

ASA
ACOUSTICAL SOCIETY
OF AMERICA

Spring 2021 Volume 17, Issue 1

An Acoustical Society of America publication

Global Positioning Systems: Over Land and Under Sea

Acoustics Today

An Acoustical Society of America publication

Spring 2021 Volume 17, Issue 1

About the Cover

Depiction of positioning of a GPS receiver using trilateration. Graphic copyright ©2021 Mark B. Weinberg, Potomac, MD. All rights reserved. More of the artist's work can be seen at www.facebook.com/lordofretirement.

ACOUSTICAL SOCIETY OF AMERICA

Publications Office
P.O. Box 809, Mashpee, MA 02649
(508) 293-1794

Follow us on Twitter @acousticsorg

Please see important *Acoustics Today* disclaimer at www.acousticstoday.org/disclaimer.

Acoustical Society of America

The Acoustical Society of America was founded in 1929 "to generate, disseminate, and promote the knowledge and practical applications of acoustics." Information about the Society can be found on the website:
www.acousticalsociety.org

Membership includes a variety of benefits, a list of which can be found at the website:
www.acousticalsociety.org/asa-membership

Acoustics Today (ISSN 1557-0215, coden ATCODK) Spring 2021, volume 17, issue 1, is published quarterly by the Acoustical Society of America, Suite 300, 1305 Walt Whitman Rd., Melville, NY 11747-4300. Periodicals Postage rates are paid at Huntington Station, NY, and additional mailing offices. POSTMASTER: Send address changes to Acoustics Today, Acoustical Society of America, Suite 300, 1305 Walt Whitman Rd., Melville, NY 11747-4300.

From the Editor

Arthur N. Popper

Our goal for *Acoustics Today* (*AT*) is that each article be interesting to, and readable by, every member of the Acoustical Society of America (ASA). Thus, I encourage everyone to take a look at each article and each "Sound Perspectives" essay in this issue. I trust most people will find something of interest and/or value in each.

The first article by Grant Eastland discusses computational methods in acoustics. Grant provides an insightful introduction to the topic and explains complex issues in ways that will help many readers appreciate that the techniques discussed could apply to their own research.

We then have a very substantial switch in topics to an article on ultrasonic hearing in non-flying terrestrial mammals. The article, written by three students, M. Charlotte Kruger, Carina Sabourin, and Alexandra Levine, and their mentor, Stephen Lomber, points out that ultrasonic hearing is actually quite common for many mammals, and that such sounds are used for communication. It is also interesting to note that this article may have more student authors than any other article in the history of *AT*. I point this out to encourage future authors to consider engaging students in articles they write for the magazine.

Our third article is by Linda Polka and Yufang Ruan. Linda and Yufang write about "baby talk." But this is not what you would immediately think of, baby language. Instead, the authors delve into the fascinating topic that a large number of ASA members are familiar with, how adults talk to babies.

The fourth article also addresses an issue that should be familiar to many (especially older) ASA members, tinnitus. Christopher Spankovich, Sarah Faucette, Celia Escabi, and Edward Lobarinas discuss this very common affliction of the auditory system and explain some of its etiology and describe how tinnitus is studied using animal models.

The fifth article by Johan Sunderg, Björn Lindblom, and Anna-Maria Hefele has another first for *AT*. Anna-Maria is not only an author but is also the subject of much of the work described, and the amazing sound files are of her special singing. Although the article focuses, to a degree, on the fascinating topic of how one singer can produce two voices at the same time, it also is a wonderful introduction to the singing voice in general.

The final article is by Lora Van Uffelen. Lora talks about global positioning systems (GPSs) and how positioning is done over land and in the water. Considering that most every reader of *AT* carries a device using GPS with them most of the time, this article provides insights into how such systems work.

This issue also has three "Sound Perspectives" essays. "Ask an Acoustician" is by Zoi-Heleni Michalopoulou. Eliza (as she is known to friends and colleagues) shares insights into her wonderful career that spans a number of ASA technical committees including Acoustical Oceanography, Signal Processing in Acoustics, and Underwater Acoustics.

The second essay is by Tyrone Porter, chair of the Committee to Improve Racial Diversity and Inclusivity (CIRDI). Tyrone introduced this committee in the December 2020 issue of *AT* (available at bit.ly/348Gbyk), and he will continue to report on this very important work in subsequent issues. In this issue, he tells us about one of the first CIRDI initiatives, working toward getting more people of color to enter the field of acoustics. As part of this article, Tyrone shares a personal story about how he became an acoustician and uses this to make the point that young people need great opportunities and great mentors to bring them into our field.

The final essay is part of what I hope will be a series over the next few years about how acoustics research is funded. These are in recognition of the fact that a significant number of ASA members pursue funding from various sources for their work, including agencies of the US government. These agencies often have compelling missions that connect to the diverse work of many of our ASA members. Thus, over the next year or two, we will invite senior leaders of these agencies to submit essays with insights about their work and passions and, where possible, information about funding opportunities. The goal is not only to share information about interesting funding organizations but perhaps also to introduce members to agencies that they might not know about but that might actually be a source of funding for their work or perhaps collaborative work with other members.

The first of these essays is by Debara L. Tucci, director of the National Institute on Deafness and Other Communication Disorders (NIDCD) at the National Institutes of Health (NIH) in Bethesda, MD. The NIH is an agency within the US Department of Health and Human Services. I invited Dr. Tucci to contribute this lead essay in part because the NIDCD provided me funding for much of my career but mostly because the NIDCD has supported many ASA members, including a number of our past presidents and our current president. I am therefore quite familiar with the NIDCD's research on hearing and speech sciences as well as the NIDCD's research on communication disorders and in the areas of taste and smell. The NIDCD has a profound impact on ASA members as the source of research and training funds to many members in animal bioacoustics, physiological and psychological acoustics, and speech acoustics.

I would also like to invite ASA members to suggest other funding agencies we might invite to provide essays. Or, if anyone reading this issue of *AT* is a funder, feel free to suggest that you do an article. The only ground rules are that the article must provide a broad overview of a program or agency that funds a large number of ASA members, perhaps across multiple technical committees. And I'd be glad to have an essay about foundations and funders outside of the United States as long as they have an impact on a substantial number of ASA members.

I also want to emphasize the interest of *AT* in having more essays on diversity and accessibility. This can be about ASA in particular or generally in relationship to STEM issues. If any member has a topic that they would like to write about, please get in touch with me. I am particularly interested in getting essays that discuss personal experiences, as Tyrone wrote in his essay in this issue, but other topics are most welcome.

Finally, I am announcing a one-time *AT* "contest." Somewhere in this issue is an advertisement that is in another language about books. The first person to find that statement and send the correct translation (as per the translation provided by the author of the statement) to the *AT* editor (apopper@umd.edu) will be mentioned in this column (along with the translation) in the summer issue and receive a small prize (a gift card) from the ASA. Good hunting!

Computational Methods and Techniques Across Acoustics

Grant C. Eastland

Sound in the World

Throughout human history, people and cultures have created sound for more than simple communication. For example, early humans likely made music using primitive flutes (Atema, 2014) and considered sound integral in the design of cities (e.g., Kolar, 2018). Furthermore, the Mayans designed structures at the ruins at Chichen Itza in Mexico that used sound for worship (Declercq et. al., 2004). Specifically, clapping in front of the stairs of the El Castillo pyramid creates a sound resembling a highly revered bird by way of a series of reflections up the stairs (available at bit.ly/3jPfOTk).

In addition to an interest in making sound, sound and vibration have also been thoroughly investigated by either empirical methods or philosophical arguments since as far back as Pythagoras (550 BCE), who applied his discoveries in mathematics to the harmonic ratios in music. He discovered that stringed instruments could be tuned, using small integer ratios of string length, so that they would consistently produce layered consonant musical intervals.

The interest and desire to study our acoustic environment continues to this day, but the methods we use have changed dramatically, and continue to change as new technologies emerge. Beginning in the seventeenth century with Robert Boyle, empirical investigation showed that sound is a vibration of conceptualized fluid particles transmitting energy from one place to another. Theoretical and empirical investigations are essential but more often require additional help to solve the problems at hand. Indeed, applying sophisticated computational methods, the basis of this article, provides a valuable tool in understanding and analyzing acoustics phenomena.

The Need for Computational Acoustics

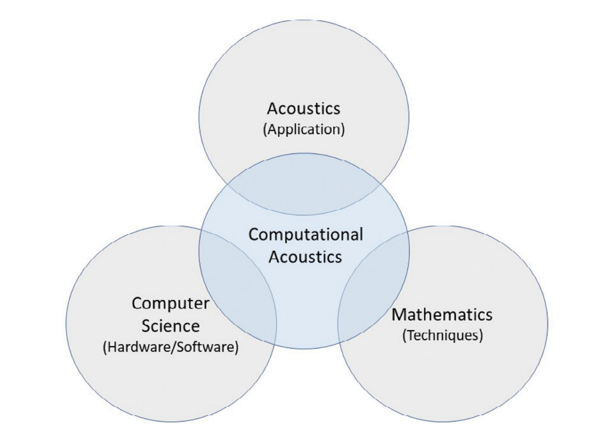
The need for computational acoustics shows itself in the difficulty in most real-world physical investigations in acoustics, often requiring solving the acoustic wave equation. Indeed, there is potential for advancement in new areas of research not contained in the traditional areas by employing computational acoustics. This is already seen from the great developments and advancements in all areas of acoustics over several decades where the complexity has required extensive use of numerical methods, optimization, computational modeling, and simulation.

Like the relationships of computational physics to mathematics and computer science, the relationship between acoustics, mathematics, and computer science define computational acoustics as described by the Venn-type diagram shown in **Figure 1**.

The Wave Equation Explained

The wave equation enables the expression of motion in a wave, and it shows itself in every area of physics including acoustics, electromagnetism, quantum mechanics, and optics, to name a few. The equation provides the

Figure 1. *Venn diagram showing the concept relationship of computational acoustics, indicating how it connects traditional acoustics with mathematics and computer science.*



 https://doi.org/10.1121/AT.2021.17.1.10

mathematical relationship between the variables of interest in acoustics, often the acoustic pressure or particle velocity and the speed of a wave. The equation relates the temporal and spatial changes to these variables, including dependence on the wavelength and frequency of the wave. As a consequence, the equation is a second-order partial differential equation of pressure or particle velocity and is three-dimensional. The pressure or particle velocity is dependent on three spatial directions and time.

In all areas of physics, solutions to problems involving the wave equation require specifying additional boundary conditions that depend on the geometry of the problem. Only in specific ideal cases with simple conditions and geometry are analytical solutions even possible. However, the wave equation is a powerful and useful tool for investigating the physics involved.

For most real problems of interest, the geometry involved is much too complicated to solve by any other means than by computational methods. For example, if one wanted to simulate the propagation of sound through the ear canal (Puria, 2020), the geometric structure would not be simple and defining real boundary conditions would make the problem too complicated to be solved any other way than by numerical solution of the wave equation.

Propagation of acoustic waves in a variety of environments is well understood and documented, but any real environment is overly complex and prediction of sound fields becomes impossible to solve analytically. For example, one may wish to determine the acoustic pressure field in a large area underwater in the ocean (e.g., Duda et al., 2019) where the environment, boundary conditions and spatial distributions of fluid properties are complicated. To solve a wave equation with such complexity, the problem is reduced to numerical solutions.

There are a bevy of techniques discussed in this article for solution of the wave equation in various situations. Several of these techniques are numerical methods applied to solve the equations directly without approximations, whereas others require a successive approximation of results.

The Emergence of Computational Methods

Since its invention in the 1930s, the digital computer has been used to solve difficult problems in physics. Early uses were in areas of nuclear physics where they performed simulations on ballistics and particle evolution for the development of the atomic bomb.

Monte Carlo Simulation

Several techniques and algorithms were developed at the Los Alamos (NM) National Laboratory by Jon von Neumann as part of his work on the atomic bomb, leading to what we now know as Monte Carlo simulations. As one might expect, Monte Carlo applications involve any phenomena that could be modeled as random or spontaneous, such as playing games of chance at a casino. Some phenomena that are modeled this way include radioactive decay and the random nature of thermal motion (Landau and Price, 1997). Additionally, Monte Carlo simulations can be used to model sound propagation in the atmosphere (Burkatovskaya et. al., 2016) where multiple scattering and the turbulent nature of the atmosphere (Blanc-Benon et al., 2002) can be taken into consideration.

Continued work in computational physics led to the discovery of chaotic behavior in nonlinear dynamics where deterministic mechanical systems exhibited seemingly random states of motion. The theoretical underpinnings of mechanics had existed for nearly half a century before computer technology made it possible to make the complicated computations needed to simulate the interactions.

Early Use of Computers in Acoustics

An early mention of using computers in acoustics is provided by a talk given at the 62nd meeting of the Acoustical Society of America by Schroeder (1961) on novel uses in room acoustics. In his abstract, Schroeder spoke of using digital computers to simulate complicated transmission of sound in rooms and simulation of spatial and frequency responses in rooms using Monte Carlo techniques. Schroeder's insight revolutionized architectural acoustics. Computational methods have proven enormously powerful in predicting acoustic performance of interior spaces and have enhanced the ability of the specialist to design spaces acoustically, such as in concert halls (Sviolja and Xiang, 2020).

The decades of improvement in computer technology and computational performance have allowed greater use of such numerical methods for acoustic wave propagation, scattering, radiation, and other acoustically related phenomena. This, in turn, has enhanced discovery and problem solving. Simulations of different phenomena have provided

Table 1. *Some relevant articles published in* Acoustics Today

Authors	Topic
Ahrens et. al., 2014	Sound field synthesis
Bruce, 2017	Speech intelligibility, signal processing
Bunting et. al., 2020	Computational acoustics
Burnett, 2015	Computer simulation of scattering
Candy, 2008	Signal processing, model-based machine learning beginnings
Duda, et. al., 2019	Ocean acoustics
Greenberg, 2018	Deep learning, languages
Hambri and Fahnline, 2007	Structural acoustics, modeling methods
Hawley et. al., 2020	Musical acoustics
Puria, 2020	Bioacoustics, hearing
Stone and Shadle, 2016	Speech production, modeling, computational fluid dynamics
Treeby, 2019	Biomedical acoustics
Vorländer, 2020	Virtual reality and music
Wage, 2018	Array signal processing and localization
Wilson et. al., 2015	Atmospheric acoustic propagation
Zurk, 2018	Underwater acoustic sensing

These papers have either a computational focus or computational relationship.

ways to investigate interactions that previously were unapproachable due to the complex nature of acoustics.

Computational acoustics, which is a combination of mathematical modeling and numerical solution algorithms, has recently emerged as a subdiscipline of acoustics. The use of approximation techniques to calculate acoustic fields with computer-based models and simulations allows for previously unapproachable problems to be solved.

The increasing computational nature of acoustics, especially in all the traditional areas, has provided a cross-disciplinary opportunity. The purpose of this paper is to show an overview of the various techniques used in computational acoustics over several of the traditional areas. I am more familiar with applications in underwater acoustics and physical acoustics, but many of the same techniques used in those areas can be applied in other areas (see **Table 1** for articles in *Acoustics Today* that discuss the use of similar techniques).

In addition, applications of machine learning (ML) that are being used in artificial intelligence research and areas of data science are also being exploited to advance research into areas including acoustic oceanography, engineering acoustics, and signal processing. This is by no means an exhaustive list, but it brings a familiarization to the areas and applications of computational acoustics and the methods found therein.

Modern Computational Methods

The numerical methods of computational acoustics are focused on taking the continuous equations and differential equations from calculus and turning them into linear algebraic equations, which are amenable to solution on digital computers. In the case of a concert hall with complex geometries that are not open to an analytic solution, computational acoustics would enable an acoustics engineer to compute a numerical solution to the wave equation to help the engineering design process, as discussed recently by Savioja and Xiang (2020).

Two of the more popular methods are the finite-difference method (FDM) and finite-element method (FEM). The FDM is a class of numerical techniques related to a general class of numerical methods known as Galerkin methods (Jensen et al., 2011; Wang et. al., 2019) that treat derivatives as algebraic differences and the continuous function in question, such as the sound field, is calculated at various points of space (Botteldooren, 1994).

For example, **Figure 2** shows how to break up the space with a grid where the sound field is calculated t as an individual element in space. Each point is calculated through iteration via a computational algorithm. The calculations are often simple enough that they could be performed with pencil and paper or a basic calculator. However, if the procedure needs to be applied to many points, there may need to be thousands to millions of computations, thereby requiring a digital computer.

In contrast to the FDM, the FEM is another numerical technique used for calculating sound fields based on dividing up a space or structure into individual elements, each of which is assumed to be constant. The space/structure is broken

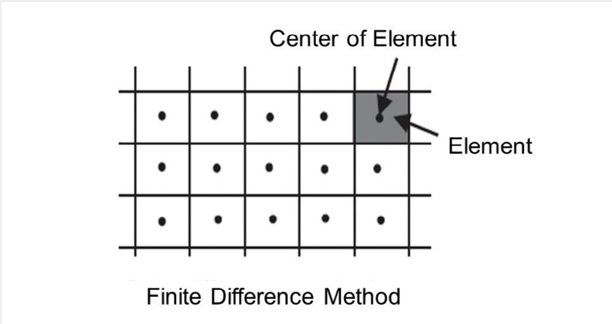


Figure 2. *Finite-difference method (FDM), a computational discrete grid concept to compute a sound field. The space is broken into a grid of boxes called elements where the sound is considered constant in each element and is summed over all space.*

up into a mesh, which looks like a wire grid applied to the structure of various shapes, often triangles. The points of the chosen mesh shape are called nodes, and these define the shape of the mesh. The goal of the method is to sum the contribution of each element to the sound field. **Figure 3** shows the conceptualization of dividing up a structure with a simple grid using a triangular mesh instead of the square boxes in **Figure 2** that divide up a structure. Although the method seems complicated, the main idea is simple.

For real-life problems, the FDM and FEM are not exclusive, and they are often applied at the same time on modern high-performance computing platforms. The FDM is simple in its application but requires some initial knowledge of conditions. The FEM is more adaptable and accurate but often requires more input data to apply.

Direct Numerical Simulation

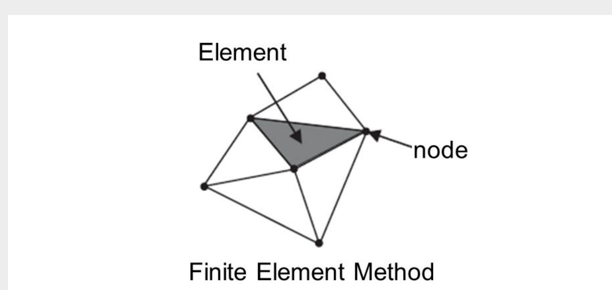
The complete mathematical treatment of complex acoustic problems in fluids begins with a set of partial differential equations known as the compressible Navier-Stokes equations. These equations describe both the flow of the fluid and the aerodynamically/hydrodynamically generated sound field. These equations are statements of conservation of momentum and mass in the fluid, describing all the dynamics.

Due to this coupling of fluid dynamics and acoustics, both fluid variables and acoustic variables may be solved directly by rewriting the equations into a form that can be fully simulated via a computer program or software package such as COMSOL or ANSYS. These types of packages are good at performing simulations of systems where multiple kinds of physics are involved, like a problem involving sound transmission through living tissue where there could be heating, density variations, and fluids in motion. Often what is required is a very precise numerical resolution due to the large changes in the length of the scales between acoustic and flow variables due to fluids in motion. The use of direct numerical simulation is often computationally challenging and is unfitting for most applications without the use of high-performance computing.

Although direct numerical simulation may be a limitation, it is often the first approach to use on a variety of problems. One such application is calculating the compressional and shear speeds of elastic waves in a material of interest utilizing measured backscattered acoustic data from a sphere made of the material. The compressional and shear speeds are related to the scattered sound in a complicated way but can be determined for spherical objects. I am not going into the complex mathematics behind the calculations; however, the method is to compute the theoretical backscattering function (Faran, 1951; Chu and Eastland, 2014). This function has discontinuities, called nulls, that are related to the compressional and shear speeds of sound in the material. The null locations and separations are dictated by these speeds.

Beginning with an initial guess of the speeds, the backscatter form function is determined. Backscattering data from the target are then matched to the form function by relating the error in the null locations and separations. Based on the selection of arbitrary nulls in the data using any nonlinear least squares method (e.g., Levenberg-Marquardt), an

Figure 3. *Finite-element method (FEM), a computational element mesh concept to compute a sound field. Each triangular division is part of the structure where the sound field can be computed assuming the triangular division is considered constant.*



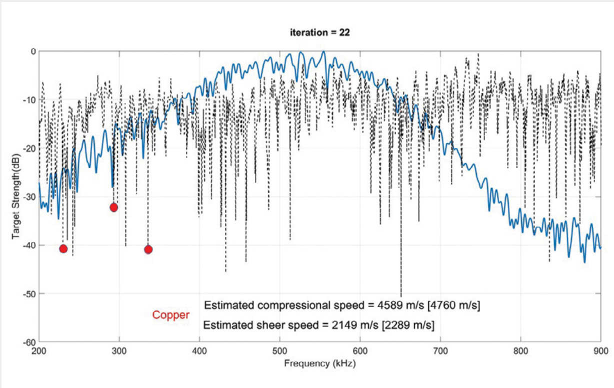


Figure 4. ***Dashed black curve**, theoretical target strength determined from the form function; **blue solid curve**, backscattered data-determined target strength of a 64-mm copper sphere, comparing theory and experiment; **red circles**, the three chosen nulls to be matched to the data, determining the compressional and shear sound speeds in the material to within at least 6% accuracy. Work was done on sonar calibration for biomass estimated acoustic surveys by the author at Northwest Fisheries Science Center (Seattle, WA).*

optimization process is selected. The goal is to minimize the error in the fitting of the data by iteratively updating the form function. The initial guess of the speeds is updated and used to recompute the form function until the desired level of error in the cost function is achieved. As one can imagine, this is a brute force method and can be computationally demanding. However, it is effective. An example of the data output is shown in **Figure 4**.

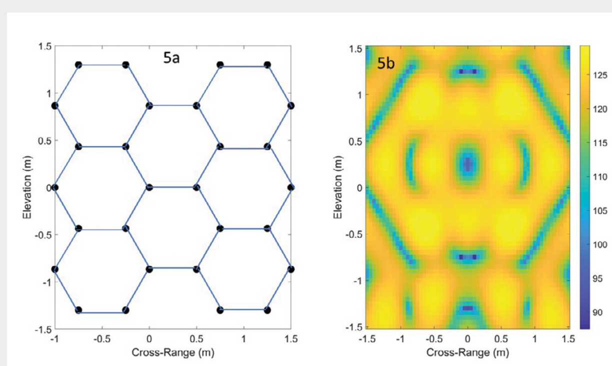
The simulation proved to be accurate in the predicted values to within 6% on the shear speed and less than 5% on the compressional speed with only 22 iterations. The computational time using MATLAB on a personal computer took nearly 10 minutes. If a higher precision is desired, the minimum error can be adjusted to get more iterations but will take much longer.

The first of these various numerical methods often used to determine the sound pressure in a computational acoustics problem is the FDM. The method takes the continuous differential equation that describes the phenomena and breaks it in to a finite algebraic set of equations. The details are left out in this article for brevity; however, the usefulness of this method is hard to deny. The method is used by breaking up the space into a grid of points, and the sound field is calculated with small changes in the field based on the nearest grid points in the space. The solution is found by solving numerically using small steps in space and time. This technique is used in multiple areas and works well for wave propagation and scattering problems.

By way of an illustration (see Bunting et. al., 2020), the application of the wave equation and discretization shows the power of computational acoustics. Assuming harmonic time dependence of pressure and applying the wave equation, one obtains the Helmholtz equation. The Helmholtz equation describes steady-state wave propagation in physics and relates to acoustic wave propagation through either the particle velocity or pressure in a fluid.

There are multiple methods utilizing a known result of the acoustic wave equation to compute the acoustic field of a sound source. A general solution for wave propagation can be written as an integral over all present sources, which are summarized as integral methods. The origin of the acoustic source must be determined a priori from some other method (e.g., a FEM simulation of a mechanical structure). The integral is taken over all sources relative to the time of the source of the signal. The sound wave arrives later at a given receiving position. Common to all integral methods is that changes in the speed of sound between the source

Figure 5. ***a:** Array of acoustic point sources arranged as several hexagonal distribution where cross-range is the lateral left/right dimension and elevation is the up/down vertical dimension as a demonstration of the method. **b:** An acoustic color plot of the simulated acoustic sound pressure level measured at a location 10 m from a source array of 31 point sources being driven in unison at 10 kHz. **Yellow** is louder than **darker blue** to a level in decibels relative to 1 μPa of acoustic pressure.*



and receiver positions cannot be justified by utilizing the theoretical solution of the wave equation.

An example application of an integral method is to calculate the acoustic field from a hexagonal array of sources treated as point sources. **Figure 5a** shows an example of an array of 31 sources arranged as a grouping of hexagons. The locations are determined computationally, with each node being treated as an acoustic point source. The field is summed over all sources, and the level is calculated at a given range. This could be done over time to create a movie of the acoustic field that can provide insight into how the acoustic wave propagates. The simulation is assumed to be underwater with a source frequency of 10 kHz. An example might be a source array, but the problem is easily simulated using integral methods. The array used in **Figure 5a** has the output given as a color plot in **Figure 5b**, respectively.

Kirchhoff Integral

Kirchhoff and Helmholtz were able to show that sound radiating from a localized source in a limited area can be described by enclosing this source area by an arbitrarily envisioned surface. The sound field inside or outside the chosen surface is calculated using the Helmholtz equation. The solution can be determined by the sum of a set of "basis" functions related to the geometry of the problem that can be used. The difficulty in the problem is determining the functions that work and is not described here due to being out of scope of this article. The calculated field on the surface directly follows from the wave equation.

A variation of the scheme allows one to calculate the pressure on the arbitrary surface using the normal particle velocity, which is the mechanism involved in acoustic transmission. The particle velocity perpendicular to the surface could be given by a FEM simulation of a moving structure. However, the modification of the method to avoid utilizing the acoustic pressure directly on the surface leads to snags, with enclosed volumes being driven at their resonant frequencies. This is a major issue in the implementation of the technique. To get around this limitation, the sound pressure is determined on the surface of the object first and then imaginary sources are added on its surface to cancel the normal particle velocity on the surface of the object.

An instance of the use of the Kirchhoff integral is to divide the physical domain into a smaller simpler set of parts for a more complex problem, which introduces the

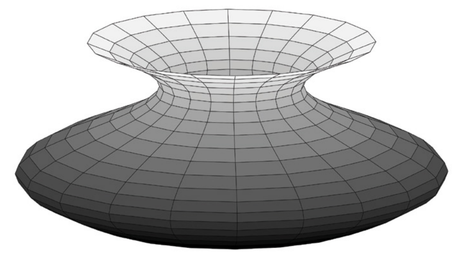


Figure 6. *Possible spatial division, called discretization, of a Helmholtz resonator in the form of a vase. Each rectangle is treated as an individual point where the sound field is considered constant.*

application of the FEM (Everstine and Henderson, 1990). This is another example of integral methods, but it solves the field by direct integration over the surface. The goal is to split the computational area into different regions so that the central acoustic equations can be solved with different sets of equations and numerical techniques.

For instance, simulating an idealized Helmholtz resonator (such as a violin or guitar) as a flower vase and solving the wave equation with boundary conditions becomes difficult due to the odd shape of the boundary. To solve this, therefore, the boundary is broken up into smaller pieces, and the acoustic field is calculated for each individual piece of the boundary. A concept figure showing what the boundary would look like is shown in **Figure 6**.

The method would then employ breaking up the vase into physical elements, as in **Figure 3**, where all the corners of the element are broken into nodes. The method just sums the acoustic field from each individual element for each node in space, assuming some constant coefficient given as $\widehat{p_i}$ for each element, approximated from the boundary and field equations. Each element would be given as some shape function given as N_i^e, which was a triangle in **Figure 3**. The total acoustic field is determined as a sum of each individual contribution such as $p(x,y,z)\approx\sum_{i=1}^{n}\widehat{p_i},N_i^e$, where x, y, and z are the obligatory spatial variables.

Machine Learning and Other Contributions

Several significant contributions have been made in different areas of investigation with the applications of computational acoustics. One of these is the incorporation of acoustic simulation methods into virtual reality systems (Vorlander, 2013, 2020). These types of systems can have real-time performance due to the advances in technology and have become paramount in the entertainment industry.

Additionally, virtual and augmented realities have been employed in training and as a diagnostic tool. In the past, there used to be latency or slowing down of simulations due to the huge amounts of data being generated. However, this is not as significant problem anymore given advances in computer technology. As a result, sound synthesis and production of indoor/outdoor surroundings can be combined with three-dimensional stereoscopic display systems through data fusion (e.g., Vorlander, 2020). The research and design applications have led to improved reality for video games and similar systems. The user experience is enhanced by adding accurately synthesized sound and allowing the listener to be able to move unrestrictedly, e.g., turn the head, to be able to perceive a more natural situation.

Moreover, the improved synthesis algorithms (e.g., Gao et al., 2020) can be used to provide more realistic conditions for psychoacoustic tests. Sound synthesis algorithms based on deterministic-stochastic signal decomposition have been applied to synthesize pitch and time scale modifications of the stochastic or random component of internal combustion engine noise (Jagla et al., 2012). The method uses a pitch-synchronous overlap-and-add algorithm, used in speech synthesis, that exploits the use of recorded engine noise data and the fact that the method does not require specific knowledge of the engine frequency. The data-based method used for speech synthesis, noise analysis, and synthesis of engine noise just mentioned is similar to what is used in ML. Applications of ML seem to have no limits in the data-driven world of today.

ML methods are based on statistics and are excellent at detecting patterns in large datasets. Applications in acoustics are fertile ground for research into ML for things such as voice recognition, source identification, and bioacoustics (e.g., Bianco et al, 2019). With technologies like Alexa or Google Home, voice recognition investigations are needed to allow the technology to work with people having different accents or pronunciations or speaking different languages. The algorithms must utilize huge datasets of recorded voices to teach the computer system to "learn" based on input. Models are developed of voices pronouncing certain common words used for searching. Variations are compared statistically to the model where the model can be improved based on additional inputs of data. The computer algorithm from the system using it essentially "learns" and incorporates that knowledge into its dataset. Although much of the research into ML and techniques are done in areas of computer science, the applications of the methods into acoustics have driven some of the more recent advances. A major method of ML, called deep learning, based on artificial neural networks that work through several layers, train systems to do everything from synthesizing music to being able to perform better than the human ear for recognition (Hawley et al., 2020).

Summary and Conclusions

The large variety of methods and applications outlined here is hardly an exhaustive depiction of computational acoustics. Due to limitations in my knowledge and the space and time to do so, only a brief introduction to the field could be given. However, hopefully, I was able to make the case for the need for the field of computational acoustics and the variety of areas of application. The uses of computational methods have driven discovery and improved understanding in a variety of areas of acoustics including sound synthesis, voice recognition, modeling of acoustic propagation, and source identification. Several techniques have been used to aid in the design of new automotive technologies by modeling the mechanical interactions of structures with different moving parts and the fluids involved.

Several of these methods are not only being used in engineering acoustics, but they are also being employed for space design for concert halls and classrooms. This type of modeling has improved noise suppression in a variety of mechanical systems. Computational techniques are being used in modeling and simulation in signal processing to utilize ML methods in the investigation of acoustic source identification and classification. The methods are being applied to areas of animal bioacoustics to aid in species identification for population monitoring, avoiding direct interaction with the animals. The methods and applications of computational acoustics are only going to grow over years to come and have become a fruitful and rewarding area of research.

Disclaimer

The opinions and assertions contained herein are my private opinions and are not to be construed as official or reflecting the views of the United States Department of Defense, specifically, the US Navy or any of its component commands.

Acknowledgments

I thank Arthur N. Popper for all his insight and feedback in developing the manuscript. I also thank my loving wife Janelle for all her support, which I would not be able to do without. Moreover, I acknowledge the love and support of my family and friends, Sabina Macy for helping me with editing, and Venus Larson, both keeping me from losing my mind.

References

Ahrens, J., Rabenstein, R., and Spors, S. (2014). Sound field synthesis for audio presentation. *Acoustics Today* 10(2), 15-25.

Atema, J. (2014). Musical origins and the Stone Age evolution of flutes. *Acoustics Today* 10(3), 25-34.

Bianco, M. J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M. A., Gannot, S., and Deledalle, C. A. (2019). Machine learning in acoustics: Theory and applications. *The Journal of the Acoustical Society of America* 146, 3590-3628. https://doi.org/10.1121/1.5133944.

Blanc-Benon, P., Lipkens, B., Dallois, L., Hamilton, M. F., and Blackstock, D. T. (2002). Propagation of finite amplitude sound through turbulence: Modeling with geometrical acoustics and the parabolic approximation. *The Journal of the Acoustical Society of America* 111, 487-498.

Botteldooren, D. (1994). Acoustical finite-difference time-domain simulation in a quasi-Cartesian grid. *The Journal of the Acoustical Society of America* 95, 2313-2319. https://doi.org/10.1121/1.409866.

Bruce, I. C. (2017). Physiologically based predictors of speech intelligibility. *Acoustics Today* 13(1), 28-35.

Bunting, G., Dohrmann, C. R., Miller, S. T., and Walsh, T. F. (2020). Solving complex acoustic problems using high-performance computations. *Acoustics Today* 16(2), 22-30. https://doi.org/10.1121/AT.2020.16.2.22.

Burkatovskaya, Y. B., Belov, V. V., Krasnenko, N. P., Shamanaeva, L. G., and Khaustov, P. A. (2016). Monte Carlo method in atmospheric acoustics. *Proceedings of the International Multi-Conference of Engineers and Computer Scientists 2016 Vol II, IMECS 2016*, Hong Kong, March 16-18, 2016.

Burnett, D. S. (2015). Computer simulation for predicting acoustic scattering from objects at the bottom of the ocean. *Acoustics Today*, 11(1), 28-36.

Candy, J. V. (2008). Signal processing in acoustics: Science or science fiction? *Acoustics Today* 4(3), 6-15.

Chu, D., and Eastland, G. C. (2014). Calibration of a broadband acoustic transducer with a standard spherical target in nearfield. *The Journal of the Acoustical Society of America* 137, 2148-2157.

Declercq, N. F., Degrieck, J., Briers, R., and Leroy, O. (2004). A theoretical study of special acoustic effects caused by the staircase of the El Castillo pyramid at the Maya ruins of Chichen-Itza in Mexico. *The Journal of the Acoustical Society of America* 116, 3328-3335. https://doi.org/10.1121/1.1764833.

Duda, T., Bonnel, J., and Heaney, K. D. (2019). Computational acoustics in oceanography: The research roles of sound field simulations. *Acoustics Today* 15(3), 28-37. https://doi.org/10.1121/AT.2019.15.3.28.

Everstine, G. C., and Henderson, F. M. (1990). Coupled finite element/boundary element approach for fluid-structure interaction. *The Journal of the Acoustical Society of America* 87, 1938-1947. https://doi.org/10.1121/1.399320.

Faran, J. J., Jr. (1951). Sound scattering by solid cylinders and spheres. *The Journal of the Acoustical Society of America* 23, 405-418.

Gao, H., Shen, Yuchen, Feng, X. and Shen, Yong (2020). Sound field synthesis of arbitrary moving sources using spectral division method. *The Journal of the Acoustical Society of America* 148, EL247-EL252. https://doi.org/10.1121/10.0001944.

Greenberg, S. (2018). Deep language learning. *Acoustics Today* 14(4), 19-27.

Hambric, S. A., and Fahnline, J. B. (2007). Structural acoustics tutorial—part 2: Sound—structure interaction. *Acoustics Today* 3(2), 9-27.

Hawley, S. H., Chatziioannou, V., and Morrison, A. (2020). Synthesis of musical instrument sounds: physics-based modeling or machine learning? *Acoustics Today*, 16(1), 20-28. https://doi.org/10.1121/AT.2020.16.1.20.

Jagla, J., Maillard, J., and Martin, N. (2012). Sample-based engine noise synthesis using an enhanced pitch-synchronous overlap-and-add method. *The Journal of the Acoustical Society of America* 132(5), 3098-3108.

Jensen, F. B., Kuperman, W. A., Porter, M. B., and Schmidt, H. (2011). *Computational Ocean Acoustics*, 2nd ed. Springer-Verlag, New York, NY.

Kolar, M. A., (2018). Archaeoacoustics: Re-sounding material culture. *Acoustics Today* 14(4), 28-37.

Landau, R. H., and Paez, M. J. (1997). *Computational Physics: Problem Solving with Computers*. John Wiley & Sons Inc., New York, NY.

Puria, S. (2020). Middle ear biomechanics: Smooth sailing. *Acoustics Today* 16(3), 27-35. https://doi.org/10.1121/AT.2020.16.3.27.

Savioja, L., and Xiang, N. (2020). Simulation-based auralization of room acoustics. *Acoustics Today* 16(4), 48-56. https://doi.org/10.1121/AT.2020.16.4.48.

Schroeder, M. R. (1961). Novel uses of digital computers in room acoustics. *The Journal of the Acoustical Society of America* 33, 1669. https://doi.org/10.1121/1.1936681.

Stone, M., and Shadle, C. H. (2016). A history of speech production research. *Acoustics Today* 12(4), 48-55.

Treeby, B. E. (2019). From biology to bytes: Predicting the path of ultrasound waves through the human body. *Acoustics Today* 15(2), 36-44. https://doi.org/10.1121/AT.2019.15.2.36.

Vorländer, M. (2013). Computer simulations in room acoustics: Concepts and uncertainties. *The Journal of the Acoustical Society of America* 133, 1203-1213.

Vorländer, M. (2020). Are virtual sounds real? *Acoustics Today* 16(1), 46-54. https://doi.org/10.1121/AT.2020.16.1.46.

Wage, K. E. (2018). When two wrongs make a right: combining aliased arrays to find sound sources. *Acoustics Today* 14(3), 48-56.

Wang, H., Sihar, I., Munoz, Raul P., and Hornikx, M. (2019). Room acoustics modeling in the time-domain with the nodal discontinuous Galerkin method. *The Journal of the Acoustical Society of America* 145(4), 2650-2663,

Wilson, D. K., Pettit, C. L., and Ostashev, V. E. (2015). Sound propagation in the atmospheric boundary layer. *Acoustics Today* 11(2), 44-53.

Zurk, L. (2018). Physics-based signal processing approaches for underwater acoustic sensing. *Acoustics Today* 14(3), 57-61.

About the Author

Grant C. Eastland
Grant.Eastland@navy.mil
Naval Undersea Warfare Center Division, Keyport
610 Dowell Street
Keyport, Washington 98345, USA

Grant C. Eastland received his PhD in physics from Washington State University (Pullman) in 2012 under the direction of Philip L. Marston, studying scattering boundary effects in acoustic imaging. He did postdoctoral research at the NOAA Northwest Fisheries Science Center in Seattle from 2012 to 2015, working on acoustic target calibration techniques. Currently, he is a physicist in test and evaluation at the Naval Undersea Warfare Center Division, Keyport (WA). His primary fields of interest are ocean acoustic propagation, scattering, and acoustic imaging. His research can be considered phenomenological, including theoretical, computational acoustic modeling, and empirical investigations.

Ultrasonic Hearing in Cats and Other Terrestrial Mammals

M. Charlotte Kruger, Carina J. Sabourin, Alexandra T. Levine, and Stephen G. Lomber

What is the first thought that comes to your mind when you read the word "ultrasound"? Most readers of *Acoustics Today* might associate ultrasound with pregnancy or perhaps specialized detection technology on ships and airplanes. Some might also think about echolocating animals. But what about terrestrial mammals? The ones that walk the earth among us? Although the use of ultrasound in echolocating mammals (e.g., bats, dolphins, and whales) is well-known, our understanding of ultrasonic perception in nonflying terrestrial mammals is limited. Here we discuss the frequencies perceived and the biological importance of ultrasound for four land-dwelling mammals as well as what is currently known about the various areas in the brain that allow these animals to process ultrasound.

What We Know About Ultrasound

Ultrasonic sounds differ from "regular" sounds because their frequencies are too high for humans to detect. The upper hearing limit for humans is considered to be 20 kHz, and sounds with a frequency above 20 kHz are considered ultrasonic. This is the agreed on definition, yet this distinction is subjectively based on the range that we, as humans, can hear and has no biological basis per se.

Despite not being able to hear ultrasound, humans often capitalize on its presence. The most familiar use would be clinical applications of ultrasound (e.g., Ketterling and Silverman, 2017). These include pregnancy scans, observation of pathology progression, and treatments such as the elimination of kidney stones (Simon et al., 2017). In industrial environments, ultrasound is used as a nondestructive test to measure the thickness and quality of objects. Even though ultrasound can be useful for humans in a variety of settings, public exposure to airborne ultrasound is suggested to also cause adverse effects, such as nausea, dizziness, and failure to concentrate (Leighton et al., 2020). However, this is not the case for many animals. Long before humans started utilizing ultrasonic frequencies, animals have been using ultrasound for various beneficial reasons.

Signals containing ultrasound play a pivotal role in the lives of many species. Well-known uses include prey detection, finding mates, and communicating with conspecifics. High frequencies have very short wavelengths and therefore attenuate more rapidly when traveling through air compared with lower frequencies. Therefore, ultrasonic production and hearing create a private communication channel that subverts detection by prey as

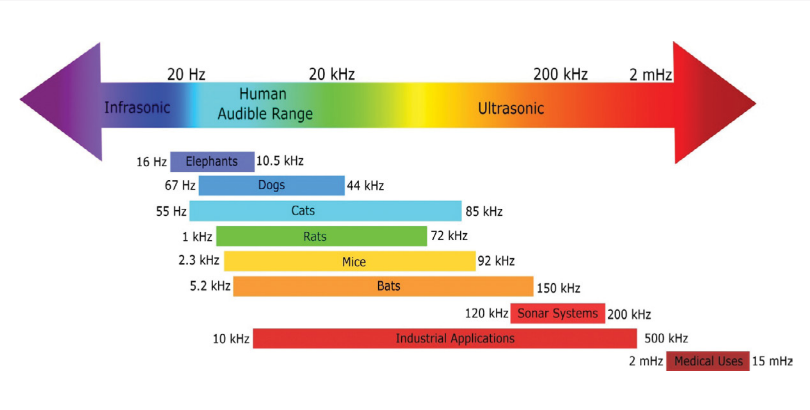


Figure 1. *Frequencies for the hearing abilities of mice and rats (*Mus musculus *and* Sigmodon hispidu*, respectively; Masterson and Heffner, 1980), elephants (*Elephas maximus*; Heffner and Heffner, 1982), domestic cats (*Felis catus*; Heffner and Heffner, 1985), domestic dogs (*Canis lupus familiaris*; Heffner, 1983), and short-tailed fruit bats (*Carollia perspicillata*; Koay et al., 2003) at 60 dB sound pressure level (SPL) as well as familiar human applications of ultrasound (Jensen, 2007; Carovac et al., 2011; Harvey et al., 2014).*

https://doi.org/10.1121/AT.2021.17.1.18

well as by predators that are unable to hear the higher frequencies (Ramsier et al., 2012). Examples of animals that can hear ultrasound include cats, dogs, bats, mice, and rats (**Figure 1**). Through technological advances, we have been able to detect, observe, study, and utilize these signals found outside our perceptual capabilities (Arch and Narins, 2008). By investigating different animals that can hear ultrasound, we better our understanding of the physiological and anatomical mechanisms behind their ability to perceive these high-frequency sounds.

The Auditory Pathway and Ultrasound

The auditory system provides animals with the ability to detect and perceive sounds over a wide range of frequencies and intensities. Sound waves travel through the outer and middle ear before being transferred to the cochlea in the inner ear. The cochlea deconstructs sounds of differing frequencies and intensities into electrical signals that can be interpreted by the brain. These electrical signals travel up the auditory pathway from the cochlea, passing through the brainstem, until eventually being relayed by the nuclei in the thalamus to their final destination, the auditory cortex.

Neurons in the auditory cortex are generally arranged according to the frequency at which they respond with the greatest sensitivity, namely their characteristic frequencies. In many animals, the characteristic frequencies of neurons progress linearly along the cortical surface as a tonotopic map (Moerel et al., 2014). This organization allows the identification of neurons responsible for conveying specific kinds of information such as ultrasound. As such, it is important to consider where these specific neurons for encoding ultrasonic frequencies are found within the cortices of terrestrial mammals and what the relevance and benefits associated with the ability to detect ultrasound might be.

Measuring the Audible Frequency Range

Audiometry experiments can provide insight into the ultrasonic abilities of different species. The point at which a sound is detected is known as the audibility threshold. As described in a previous issue of *Acoustics Today* (Dent, 2017), psychophysical approaches are often employed to measure perceptual thresholds in nonhuman species. Psychophysical approaches encapsulate experimental designs where a physical stimulus is presented to a subject and the neural and/or sensory responses evoked by the stimulus are examined. Psychoacoustics, one form of psychophysics, analyzes the relationship between auditory stimuli and neural events by employing various conditioning techniques (Dent, 2017). The results of the different conditions tested are often depicted using an audiogram (**Figure 2**).

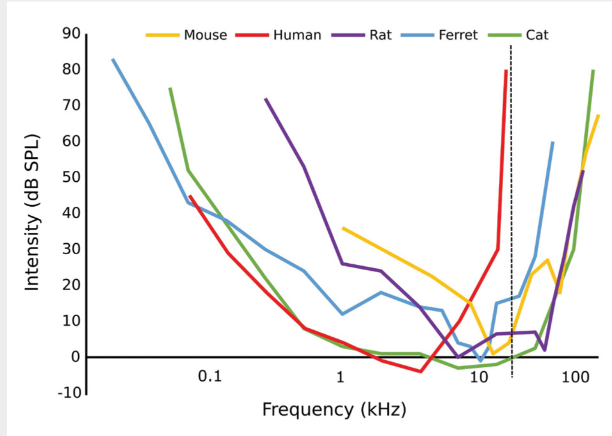


Figure 2. *A behavioral audiogram for the mouse (*Mus musculus*; Ehret, 1974), ferret (*Mustela putorius furo*; Kelly et al., 1986), human (*Homo sapiens*; Sivian and White, 1933), rat (Kelly and Masterson, 1977), and cat (*Felis catus*; Heffner and Heffner, 1985).* ***Dashed vertical line****, beginning of the ultrasonic range (20 kHz). Data represent the lowest sound level detected at each frequency.*

Biological Importance

Rodents

Small rodents such as rats and mice emit and perceive ultrasonic sounds to communicate with conspecifics for a variety of social communicative interactions, including courtship and mating, aggression and territoriality, reproduction, and to alert conspecifics (Arch and Narins, 2008).

Male mice produce ultrasonic vocalizations between 48 and 79 kHz in response to female pheromones to attract them as a potential mate (Gourbal et al., 2004) and emit vocalizations up to 75 kHz when sniffing or mounting female mice (Matsumoto and Okanoya, 2016). Furthermore, mouse pups produce isolation calls with frequencies between 50 and 80 kHz when they are separated from their nest (Hofer et al., 2001). Because mice can hear frequencies between 1.5 and 92 kHz at 60 dB sound pressure level (SPL; Heffner et al., 2001), the pup vocalizations reliably elicit a retrieval response from the mothers (Dunlap and Liu, 2018). The frequencies used in courtship vocalizations produced by male mice to attract females shows the sex-specific relevance of ultrasound production and hearing.

Similar to mice, adult rats have two main purposes for emitting ultrasonic vocalizations as a form of communication: alarm calls at 22 kHz to warn conspecifics of danger and calls at 50 kHz for social cooperation and affiliative behavior (Wright et al., 2010). Rats generally emit vocalizations with frequencies that fall within their hearing range (between 250 Hz and 80 kHz). For example, infant rats can emit vocalizations between 40 and 65 kHz when they are separated from their nest, and adult rats can emit ultrasonic calls to solicit sexual behavior from the opposite sex (Portfors, 2007). On hearing the 50 kHz vocalizations from male rats, females display a series of attracting behaviors, increasing the likelihood of the male approaching and copulating (Portfors, 2007). Rodents therefore rely on ultrasound for their survival whether it is for communicating with conspecifics, attracting mates, or evading predators.

Carnivores

Unlike rodents, there are only limited data available on the evolution and biological importance of ultrasonic hearing in carnivores. Carnivores, aside from carnivorous rodents like the northern grasshopper mouse (*Onychomys leucogaster*), are seldom known to produce or use ultrasonic frequencies for communication (Brown et al. 1978; Farley et al., 1987). Even so, many carnivores can perceive sounds with ultrasonic frequencies. It is thought that perhaps, at one point in history, the common ancestor of carnivores used ultrasound for prey detection (Heffner and Heffner, 1985; Kelly et al., 1986). However, as discussed in ***Rodents***, prey (such as mice or rats) primarily communicate at frequencies above the hearing range of carnivores (Kelly and Masterton, 1977).

Phillips and colleagues (1988) determined that ferrets (*Mustela putorius furo*) can detect sounds from 40 Hz to approximately 40 kHz. Ferrets provide a useful model for investigating the development, organization, and plasticity of the auditory cortex because the onset of hearing in ferrets occurs late compared with other mammals (Moore, 1982). Before their ear canals open, newborn ferrets, known as kits, produce high-frequency vocalizations often above 16 kHz. Lactating female ferrets respond to these kit vocalizations (Shimbo, 1992) similar to the rodent behavior described in ***Rodents***. Overall, ferrets provide useful models for investigating different aspects of hearing and hearing loss, given that their hearing range largely overlaps that of humans (Fritz et al., 2007).

Another common carnivore model used for auditory research is the domestic cat (*felis catus*). The sensitive hearing range of cats is commonly believed to be between 5 and 32 kHz, although there are notable discrepancies in the literature regarding their hearing range limits (**Figure 3**). The literature agrees that cats can hear ultrasonic frequencies, but the full extent of their perception remains unclear. The lower limit of hearing is generally reported as approximately 125 Hz, but the upper limit is not well defined.

Most sources report the upper limit as the maximum frequency tested. As such, the upper hearing limit of cats is not commonly described as greater than 60 kHz (**Figure 3**), and, in some cases, the reported upper limit corresponds to the highest frequency of sound tested in the respective study. This is true for both electrical stimulation experiments, where electrical impulses are applied to neurons in the auditory pathway, and behavioral experiments. One exception is a study by Heffner and Heffner (1985) who tested frequencies up to 92 kHz and reported the upper hearing limit as 85 kHz. Therefore, it is possible

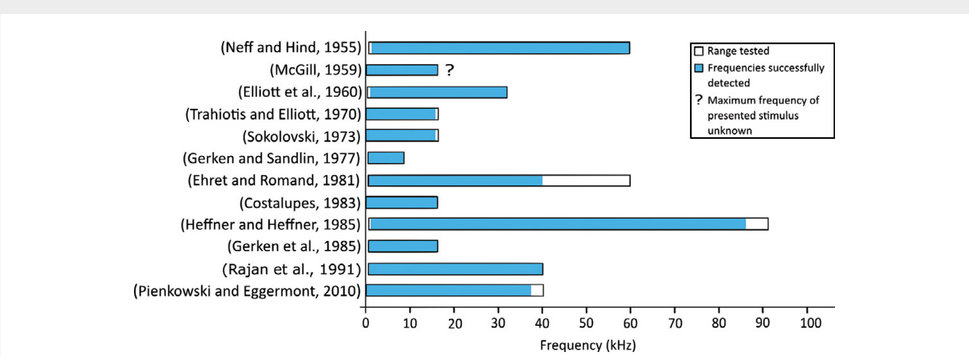
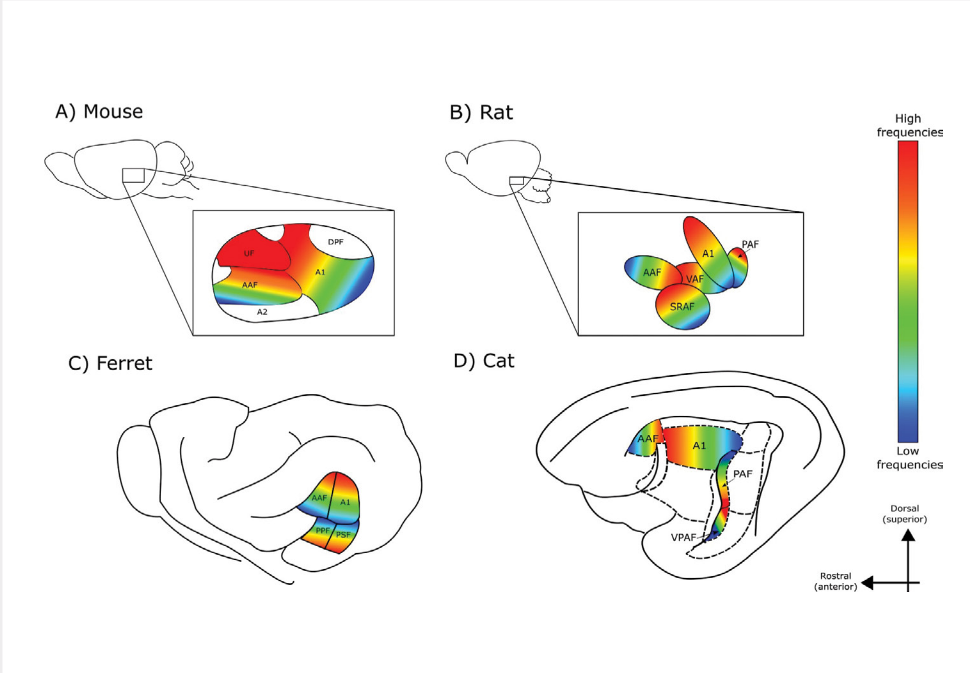


Figure 3. *Various reported hearing ranges for cats compared with the range of frequencies of the presented stimuli in each respective study, all of which are cited in* **References**.

Figure 4. *The tonotopic organization of the auditory cortex of four mammals.* ***A:*** *mouse;* ***B:*** *rat;* ***C:*** *ferret;* ***D:*** *cat. The brains are shown in the sagittal plane and show the primary auditory field (A1), secondary auditory field (A2), anterior auditory field (AAF), dorsoposterior field (DPF), ultrasonic field (UF), posterior auditory field (PAF), ventral auditory field (VAF), suprarhinal auditory field (SRAF), posterior pseudosylvian field (PPF), posterior suprasylvian field (PSF), and ventroposterior auditory field (VPAF).*



that the upper hearing limit of cats exceeds 60 kHz and that there could be neurons present in the cortex specialized for these ultrasonic frequencies.

Cortical Representation of Ultrasonic Frequencies

Mice, rats, ferrets, and cats are commonly used as animal models for acoustic research. The biological importance of ultrasound to these mammals is further reflected by the allotment of cortical space for ultrasonic sound perception in their respective auditory cortices. As such, it is crucial to validate as well as expand our current understanding of their hearing abilities, especially the neural correlates underlying the perception of ultrasonic frequencies.

Mice

In the mouse brain (**Figure 4A**), five auditory cortical fields can be delineated in both hemispheres: primary auditory field (A1), anterior auditory field (AAF), secondary auditory field (A2), dorsoposterior field (DPF), and ultrasonic field (UF) (Stiebler et al., 1997). The A1 and AAF regions are both tonotopically organized but with reverse gradients. The properties of the neurons within these two fields are similar. For example, the frequency ranges for neurons found in both the A1 and AAF are between 2 and 45 kHz.

The mouse was the first animal where a specialized cortical region for processing ultrasonic frequencies was identified (Hofstetter and Ehret, 1992). Frequencies between 40 and 70 kHz are represented in the UF, with approximately 50% of neurons responding to frequencies between 50 and 60 kHz. However, unlike the A1 and AAF, the UF is not tonotopically organized (Stiebler et al., 1997), and it is still not clear whether the UF should be considered a part of the primary auditory fields alongside the A1 and AAF.

Tsukano and colleagues (2015) showed that the dorsomedial field (DM), previously thought to be part of dorsal A1, is a separate area specialized for ultrasonic perception. This region contains neurons highly responsive to vocalizations, with frequencies above 40 kHz, demonstrating how certain neurons in mouse cortex respond best to frequencies of behaviorally relevant sound features. This type of cortical organization can also be seen in other rodents that rely on ultrasound for survival.

Rats

The central auditory system of rats is comparable to that of mice in both anatomical and functional organization. Five distinct cortical fields have been identified in the rat brain, and high-frequency neurons can be found in the following regions: A1, AAF, posterior auditory field (PAF), ventral auditory field (VAF), and suprarhinal auditory field (SRAF).

In terms of tonotopic organization, the A1 shows a progression of characteristic frequencies from low (~1 kHz) to high (~60 kHz) along a posterior-to-anterior gradient (Polley et al., 2007). The tonotopic gradient then reverses in a mirror-like fashion at the posterior and anterior borders of the A1 to form the boundaries of the PAF and AAF, respectively (**Figure 4B**) (Rutkowski et al., 2003; Polley et al., 2007).

Unlike mice, an ultrasonic field has not been identified in rats, although, because the tonotopic organization in the rat is comparable to that of the mouse, Kalatsky and colleagues (2005) hypothesized that a distinct region representing ultrasonic frequencies might likely also be present in rats. Overall, despite the similarities between the suggested cortical maps, further investigation is needed to improve our understanding of ultrasonic representations in the auditory cortex of rodents. This could potentially lead to discoveries that could, in turn, be extended to other mammals.

Ferrets

Like the auditory cortex of other mammals, the ferret auditory cortex is divided into multiple subregions. These include the two primary areas, the A1 and AAF (Bajo et al., 2006) and the secondary areas: anterior dorsal field, posterior pseudosylvian field (PPF), and posterior suprasylvian field (PSF). The PPF and the PSF are found immediately ventral to the A1 (**Figure 4C**). Bizley and colleagues (2005) described the functional organization of the different regions within the ferret auditory cortex and subsequently mapped the tonotopic organization of these areas.

As discussed for mice and rats and also for most other mammals, the frequencies in these fields are organized from high to low in a rostrocaudal manner, with frequency reversals taking place at the borders between adjacent fields (Bizley et al., 2005). However, this reversal pattern is not present in ferrets. Instead, the frequencies are organized where the gradients of the A1 and AAF meet dorsally and decrease ventrally (**Figure 4C**) (Kaas, 2011). Therefore, the A1 and AAF are organized tonotopically, with higher frequencies represented toward the dorsal tip. The physiological properties of the ferret A1 (such as tonotopic organization and neuronal properties) are similar to those seen in the cat A1 (Kaas, 2011), but when comparing audiograms of ferrets and cats, the ferret's audiogram is shifted toward lower frequencies (**Figure 2**).

Cats

Similar to ferrets, the cat auditory cortex can be divided into one or more primary areas and several secondary areas (Bizley et al., 2005). To help describe the functional and tonotopic organization of the cat auditory cortex, Reale and Imig (1980) analyzed how clusters of neurons (and sometimes single neurons) respond to various frequencies. In addition to describing the tonotopic organization of the core auditory region, the A1 and AAF, Reale and Imig (1980) also described the presence and tonotopic organization of the PAF and the ventroposterior auditory field (VPAF). Furthermore, they delineated the belt auditory region into the A2, temporal area (T), dorsoposterior area (DP), and ventral area (V).

More recently, Hall and Lomber (2015) confirmed the four functionally distinct tonotopic areas within the cat auditory cortices (A1, AAF, PAF, and VPAF) and reported a reversal in tonotopic gradients between neighboring regions (**Figure 4D**). In the cat, the A1 increases in its tonotopic gradient as it extends from the anterior division of the posterior ectosylvian sulcus (PES) to the posterior portion of the anterior ectosylvian sulcus (AES). At the posterior edge of the PES, the A1 reaches the minimal values of its tonotopic gradient, forming a low-frequency reversal border as it nears the PAF (Hall and Lomber, 2015). High-frequency reversal borders also exist at the A1-AAF and PAF-VPAF borders and are likely a location where ultrasonic selective neurons may be found.

Ultrasonic-Selective Neurons in the Cat Auditory Cortex

Following the principles of tonotopic organization, it seems that neurons with the highest characteristic frequencies could potentially be located at the periphery of each auditory region. Neurons can be classified as being either broadly or narrowly tuned, responding maximally to a large range or a narrow range of frequencies, respectively. This classification provides insight into the type of sensory input the neurons convey and their roles within a specific cortical field. High-frequency selective neurons have been found to be narrowly tuned (Phillips and Irvine, 1982), supporting the claim that high-frequency reversal borders (e.g., between the A1 and AAF) contain mostly such neurons. For example, Carrasco and Lomber (2010) identified neurons selective for frequencies reaching 60 kHz around the border between the A1 and AAF.

However, it is possible that the high-frequency reversal borders also contain broadly tuned subpopulations of neurons, functioning to integrate a wide range of ultrasonic sensory input for both the A1 and AAF. Researchers sometimes avoid probing these regions near the borders to prevent "contamination" from the accidental recording of neurons from neighboring regions (Carrasco et al., 2015). Due to this precaution, studies may fail to record from the neurons with the highest characteristic frequencies, limiting our understanding of the cortical representation of ultrasonic stimuli in cats and other terrestrial mammals.

Conclusion

Ultrasound is essential to the lives of many animals, evidenced by the magnitude of cortical space allocated specifically for ultrasound in some species. Despite the biological importance of ultrasound, researchers might have been underestimating the ultrasonic hearing abilities of many terrestrial mammals. It is therefore also possible that the extent of the ultrasonic abilities of some terrestrial mammals, as discussed in this article, has not yet been conclusively established. Until this matter is clarified, the location in the brain where these frequencies are encoded also remains uncertain. Further investigations may elucidate uncertainties in our understanding of the role of ultrasonic frequencies in auditory neuroscience as a whole.

References

Arch, V. S., and Narins, P. M. (2008). "Silent" signals: Selective forces acting on ultrasonic communication systems in terrestrial vertebrates. *Animal Behaviour* 76, 1423-1428. https://doi.org/10.1016/j.anbehav.2008.05.012.

Bajo, V. M., Nodal, F. R., Bizley, J. K., Moore, D. R., and King, A. J. (2006). The ferret auditory cortex: descending projections to the inferior colliculus. *Cerebral Cortex* 17, 475-491. https://doi.org/10.1093/cercor/bhj164.

Bizley, J. K., Nodal, F. R., Nelken, I., and King, A. J. (2005). Functional organization of ferret auditory cortex. *Cerebral Cortex* 15, 1637-1653. https://doi.org/10.1093/cercor/bhi042.

Brown, K. A., Buchwald, J. S., Johnson, J. R., and Mikolich, D. J. (1978). Vocalization in the cat and kitten. *Developmental Psychobiology* 11, 559-570. https://doi.org/10.1002/dev.420110605.

Carovac, A., Smajlovic, F., and Junuzovic, D. (2011). Application of ultrasound in medicine. *Acta Informatica Medica* 19(3), 168-171. https://doi.org/10.5455/aim.2011.19.168-171.

Carrasco, A., and Lomber, S. G. (2010). Reciprocal modulatory influences between tonotopic and nontonotopic cortical fields in the cat. *Journal of Neuroscience* 30, 1476-1487. https://doi.org/10.1523/JNEUROSCI.5708-09.2009.

Carrasco, A., Kok, M. A., and Lomber, S. G. (2015). Effects of core auditory cortex deactivation on neuronal response to simple and complex acoustic signals in the contralateral anterior auditory field. *Cerebral Cortex* 25, 84-96. https://doi.org/10.1093/cercor/bht205.

Costalupes, J. A. (1983). Temporal integration of pure tones in the cat. *Hearing Research* 9, 43-54. https://doi.org/10.1016/0378-5955(83)90133-8.

Dent, M. (2017). Animal psychoacoustics. *Acoustics Today* 13(3), 19-26.

Dunlap, A. G., and Liu, R. C. (2018). Rodent ultrasonic vocalization as a window into auditory cortical processing and plasticity. In S. M. Brudzynski (Ed.), *Handbook of Behavioral Neuroscience*. Academic Press, London, UK, pp. 83-93.

Ehret, G. (1974). Age-dependent hearing loss in normal hearing mice. *Naturwissenschaften* 61, 506-507. https://doi.org/10.1007/BF00622976.

Ehret, G., and Romand, R. (1981). Postnatal development of absolute auditory thresholds in kittens. *Journal of Comparative and Physiological Psychology* 95, 304-311. https://doi.org/10.1037/h0077770.

Elliott, D. N., Stein, L., and Harrison, M. J. (1960). Determination of absolute-intensity thresholds and frequency-difference thresholds in cats. *Journal of the Acoustical Society of America* 32, 380-384. https://doi.org/10.1121/1.1908071.

Farley, S. D., Lehner, P. N., Clark, T., and Trost, C. (1987). Vocalizations of the Siberian ferret (*Mustela eversmanni*) and comparisons with other mustelids. *Journal of Mammalogy* 68, 413-416. https://doi.org/10.2307/1381487.

Fritz, J. B., Elhilali, M., David, S.V., and Shamma, S. A. (2007). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hearing Research* 229(1-2), 186-203. https://doi.org/10.1016/j.heares.2007.01.009.

Gerken, G. M., and Sandlin, D. (1977). Auditory reaction time and absolute threshold in cat. *Journal of the Acoustical Society of America* 61, 602-606. https://doi.org/10.1121/1.381306.

Gerken, G. M., Saunders, S. S., Simhadri-Sumithra, R., and Bhat, K. H. V. (1985). Behavioral thresholds for electrical stimulation applied to auditory brainstem nuclei in cat are altered by injurious and noninjurious sound. *Hearing Research* 20, 221-231. https://doi.org/10.1016/0378-5955(85)90027-9.

Gourbal, B. E. F., Barthelemy, M., Petit, G., and Gabrion, C. (2004). Spectrographic analysis of the ultrasonic vocalisations of adult male and female BALB/c mice. *Naturwissenschaften*, 91(8), 381-385. https://doi.org/10.1007/s00114-004-0543-7.

Hall, A. J., and Lomber, S. G. (2015). High-field fMRI reveals tonotopically-organized and core auditory cortex in the cat. *Hearing Research* 325, 1-11. https://doi.org/10.1016/j.heares.2015.03.003.

Harvey, G., Gachagan, A., and Mutasa, T. (2014). Review of high-power ultrasound-industrial applications and measurement methods. *IEEE Transactions on Ultrasonics Ferroelectrics, and Frequency Control*, 61(3), 481-495. https://doi.org/10.1109/TUFFC.2014.2932.

Heffner, H. E. (1983). Hearing in large and small dogs: Absolute thresholds and size of the tympanic membrane. *Behavioral Neuroscience* 97(2), 310-318. https://doi.org/10.1037/0735-7044.97.2.310.

Heffner, R. S., and Heffner, H. E. (1982). Hearing in the elephant (*Elephas maximus*): Absolute sensitivity, frequency discrimination, and sound localization. *Journal of Comparative and Physiological Psychology* 96(6), 926-944. https://doi.org/10.1037/0735-7036.96.6.926.

Heffner, R. S., and Heffner, H. E. (1985). Hearing range of the domestic cat. *Hearing Research* 19, 85-88. https://doi.org/10.1016/0378-5955(85)90100-5.

Heffner, R. S., Koay, G., and Heffner, H. E. (2001). Audiograms of five species of rodents: Implications for the evolution of hearing and the perception of pitch. *Hearing Research* 157(1-2), 138-152. https://doi.org/10.1016/S0378-5955(01)00298-2.

Hofer, M. A., Shair, H. N., and Brunelli, S. A. (2001). Ultrasonic vocalizations in rat and mouse pups. *Current Protocols in Neuroscience* 17(1), 8.14.1-8.14.16. https://doi.org/10.1002/0471142301.ns0814s17.

Hofstetter, K. M. and Ehret, G. (1992). The auditory cortex of the mouse: connections of the ultrasonic field. *Journal of Comparative Neurology* 323, 370-386. https://doi.org/10.1002/cne.903230306.

Jensen, J. A. (2007). Medical ultrasound imaging. *Progress in Biophysics and Molecular Biology* 93(1-3), 153-165. https://doi.org/10.1016/j.pbiomolbio.2006.07.025.

Kaas, J. H. (2011). The evolution of auditory cortex: The core areas. In J. Winer Jand C. Schreiner. (Eds.), *Auditory Cortex*. Springer US, New York, NY, pp. 407-427. https://doi.org/10.1007/978-1-4419-0074-6_19.

Kalatsky, V. A., Polley, D. B., Merzenich, M. M., Schreiner, C. E., and Stryker, M. P. (2005). Fine functional organization of auditory cortex revealed by Fourier optical imaging. *Proceedings of the National Academy of Sciences of the United States of America* 102(37), 13325-13330. https://doi.org/10.1073/PNAS.0505592102.

Kelly, J. B., and Masterton, B. (1977). Auditory sensitivity of the albino rat. *Journal of Comparative and Physiological Psychology* 91, 930-936.

Kelly, J. B., Kavanagh, G. L., and Dalton, J. C. H. (1986). Hearing in the ferret (*Mustela putorius*): Thresholds for pure tone detection. *Hearing Research* 24, 269-275. https://doi.org/10.1016/0378-5955(86)90025-0.

Ketterling, J. A., and Silverman, R. H. (2017). Clinical and preclinical applications of high-frequency ultrasound. *Acoustics Today* 13(1), 41-51.

Koay, G., Heffner, R. S., Bitter, K. S., and Heffner, H. E. (2003). Hearing in American leaf-nosed bats. II: *Carollia perspicillata*. *Hearing Research* 178(1-2), 27-34. https://doi.org/10.1016/S0378-5955(03)00025-X.

Leighton, T. G., Lineton, B., Dolder, C., and Fletcher, M. D. (2020). Public exposure to airborne ultrasound and very high frequency sound. *Acoustics Today* 16(3), 17-25.

Masterton, B., and Heffner, H. (1980). Hearing in Glires: Domestic rabbit, cotton rat, feral house mouse, and kangaroo rat. *The Journal of the Acoustical Society of America* 68(6), 1584–1599. https://doi.org/10.1121/1.385213.

Matsumoto, Y. K., and Okanoya, K. (2016). Phase-specific vocalizations of male mice at the initial encounter during the courtship sequence. *PLoS ONE* 11, e0147102. https://doi.org/10.1371/journal.pone.0147102.

McGill, T. E. (1959) Auditory sensitivity and the magnitude of cochlear potentials. *The Annals of Otology, Rhinology and Laryngology* 68, 193-207.

Moerel, M., De Martino, F., and Formisano, E. (2014). An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience* 8, 225. https://doi.org/10.3389/fnins.2014.00225.

Moore, D. R. (1982). Late onset of hearing in the ferret. *Brain Research* 253(1-2), 309-311. https://doi.org/10.1016/0006-8993(82)90698-9.

Neff, W. D., and Hind, J. E. (1955). Auditory thresholds of the cat. *Journal of the Acoustical Society of America* 27, 480-483. https://doi.org/10.1121/1.1907941.

Phillips, D. P., and Irvine, D. R. F. (1982). Properties of single neurons in the anterior auditory field (AAF) of cat cerebral cortex. *Brain Research* 24, 237-244. https://doi.org/10.1016/0006-8993(82)90581-9.

Phillips, D. P., Judge, P. W., and Kelly, J. B. (1988). Primary auditory cortex in the ferret (*Mustela putorius*): Neural response properties and topographic organization. *Brain Research* 443, 281-294. https://doi.org/10.1016/0006-8993(88)91622-8.

Pienkowski, M., and Eggermont, J. J. (2010). Intermittent exposure with moderate-level sound impairs central auditory function of mature animals without concomitant hearing loss. *Hearing Research* 261(1-2), 30-35. https://doi.org/10.1016/j.heares.2009.12.025.

Polley, D. B., Read, H. L., Storace, D. A., and Merzenich, M. M. (2007). Multiparametric auditory receptive field organization across five cortical fields in the albino rat. *Journal of Neurophysiology* 97, 3621-3638. https://doi.org/10.1152/jn.01298.2006.

Portfors, C. V. (2007). Types and functions of ultrasonic vocalizations in laboratory rats and mice. *Journal of the American Association for Laboratory Animal Science* 46, 28-34.

Rajan, R., Irvine, D. R. F., and Cassell, J. F. (1991). Normative N1 audiogram data for the barbiturate-anaesthetised domestic cat. *Hearing Research* 53, 153-158. https://doi.org/10.1016/0378-5955(91)90222-U.

Ramsier, M. A., Cunningham, A. J., Moritz, G. L., Finneran, J. J., Williams, C. V., Ong, P. S., Gursky-Doyen, S. L., and Dominy, N. J. (2012). Primate communication in the pure ultrasound. *Biology Letters* 8, 508-511. https://doi.org/10.1098/rsbl.2011.1149.

Reale, R. A., and Imig, T. J. (1980). Tonotopic organization in auditory cortex of the cat. *Journal of Comparative Neurology* 192, 265-291. https://doi.org/10.1002/cne.901920207.

Rutkowski, R. G., Miasnikov, A. A., and Weinberger, N. M. (2003). Characterisation of multiple physiological fields within the anatomical core of rat auditory cortex. *Hearing Research* 181, 116-130. https://doi.org/10.1016/S0378-5955(03)00182-5.

Shimbo, F. M. (1992), *A Tao Full of Detours, the Behavior of the Domestic Ferret*. Ministry of Publications, Elon College, NC.

Simon, J. C., Maxwell, A. D., and Bailey, M. R. (2017). Some work on the diagnosis and management of kidney stones with ultrasound. *Acoustics Today* 13(4), 52-59.

Sivian, L. J., and White, S. D. (1933). On minimum audible sound fields. *Journal of the Acoustical Society of America* 4, 288-321. https://doi.org/10.1121/1.1915608.

Sokolovski, A. (1973). Normal threshold of hearing for cat for free-field listening. *Archiv Für Klinische Und Experimentelle Ohren-, Nasen- Und Kehlkopfheilkunde* 203(3), 232-240. https://doi.org/10.1007/BF00344934.

Stiebler, I., Neulist, R., Fichtel, I., and Ehret, G. (1997). The auditory cortex of the house mouse: Left-right differences, tonotopic organization and quantitative analysis of frequency representation. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology* 181, 559-571. https://doi.org/10.1007/s003590050140.

Trahiotis, C., and Elliot, D. M. (1970). Behavioral investigation of some possible effects of sectioning the crossed olivocochlear bundle. *Journal of the Acoustical Society of America* 47, 592-596. https://doi.org/10.1121/1.1911934.

Tsukano, H., Horie, M., Bo, T., Uchimura, A., Hishida, R., Kudoh, M., Takahashi, K., Takebayashi, H., and Shibuki, K. (2015). Delineation of a frequency-organized region isolated from the mouse primary auditory cortex. *Journal of Neurophysiology* 113, 2900-2920. https://doi.org/10.1152/jn.00932.2014.

Wright, J. M., Gourdon, J. C., and Clarke, P. B. S. (2010). Identification of multiple call categories within the rich repertoire of adult rat 50-kHz ultrasonic vocalizations: Effects of amphetamine and social context. *Psychopharmacology* 211, 1-13. https://doi.org/10.1007/s00213-010-1859-y.

About the Authors

M. Charlotte Kruger
charlotte.kruger@mail.mcgill.ca

Department of Physiology
McGill University
Montréal, Québec H3G 1Y6, Canada

M. Charlotte Kruger graduated from the University of Western Ontario (London, ON, Canada) with a BSc Honors Specialization in Biology in 2019. She is a current graduate student in the Cerebral Systems Laboratory at McGill University (Montréal, QC, Canada). She is investigating the ultrasonic hearing abilities of the cat and the location where ultrasonic frequencies might be encoded in the cat brain, to better understand the role of ultrasonic hearing in auditory neuroscience.

Carina J. Sabourin
carina.sabourin@mail.mcgill.ca
Department of Physiology
McGill University
Montréal, Québec H3G 1Y6, Canada

Carina J. Sabourin graduated from Queen's University (Kingston, ON, Canada) with BSc Honors Specialization in Life Sciences. She recently joined the Cerebral Systems Laboratory at McGill University (Montreal, QC, Canada) as a graduate student pursuing an M.Eng in Biological and Biomedical Engineering. She is using awake electrophysiological recording techniques to investigate the auditory cortex in animal models. She is currently using system identification methods to model the temporal modulation of auditory neurons in response to biologically important sounds.

Alexandra T. Levine
alevin23@uwo.ca
Department of Psychology
Western University
London, Ontario N6A 3K7, Canada

Alexandra T. Levine is a postdoctoral associate in the Neuroplacity in Sensory Systems Lab (NiSSL) at Western University (London, ON, Canada) researching brain plasticity following sensory loss. She is currently working on a range of projects involving ultrahigh-field magnetic resonance imaging (MRI), including structural and functional differences in the brain resulting from hearing loss and neural correlates of behavioral advantages seen in such individuals. She received her BSc in psychology and cognitive neuroscience from The University of Nottingham (University Park, Nottingham, UK) and her MSc in cognitive neuroscience and PhD in cognitive neuroscience and neuroimaging from the University of York (UK).

Stephen G. Lomber
steve.lomber@mcgill.ca
Department of Physiology
McGill University
Montréal, Québec H3G 1Y6, Canada

Stephen G. Lomber is a professor of physiology at McGill University (Montréal, QC, Canada) where he holds the Canada Research Chair in Brain Plasticity and Development. Dr. Lomber received his BSc from the University of Rochester (NY) and his PhD from the Boston University School of Medicine (MA), both in neuroscience. Dr. Lomber's laboratory (www.cerebralsystems.ca) uses an integrated approach of psychophysics, electrophysiological recording, neuroanatomical approaches, and functional imaging to examine multisensory processing in auditory cortex. Work in the laboratory examines cortical plasticity in the presence and absence of acoustic input and following the initiation of auditory processing through the use of cochlear prosthetics.

The Ins and Outs of Baby Talk

Linda Polka and Yufang Ruan

It is usually no secret when there is a baby in the room. Infants attract our attention, and we immediately and instinctively change our speech when we engage with them. "Baby talk" fills the air. This distinct speech register, also known as *motherese* or more formally as *infant-directed speech* (IDS), has been observed across diverse languages and cultures. Babies demonstrate a clear preference for IDS. The strong endorsement of IDS by infants continues to fuel the curiosity of scientists, clinicians, and caregivers about this common speech form and how it shapes infant development.

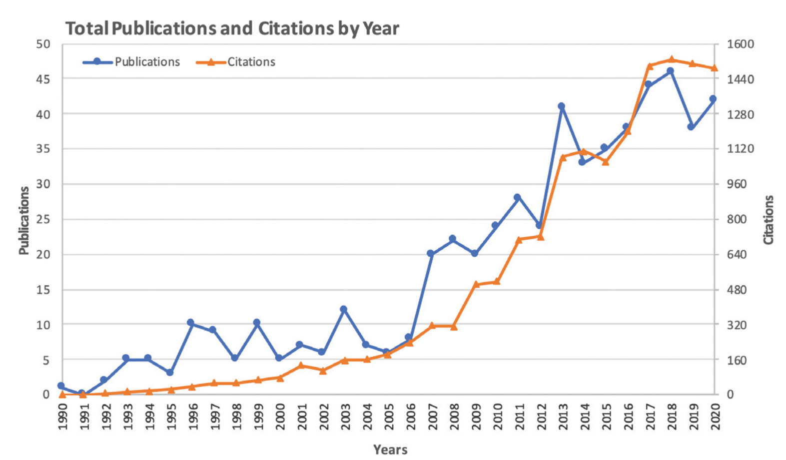
In the research world, "infant" is often defined as a child under 2 years of age. In recent years, scientific interest in IDS has increased dramatically. **Figure 1** shows that the number of publications on the topic of IDS and citations of this work increased markedly since 2006. In 2019, the Acoustical Society of America sponsored two well-attended special sessions devoted to IDS (see acousticstoday.org/ASA177TuesAbstracts, pages 1728-1731 and 1763-1767). IDS is undeniably a hot topic.

Although producing IDS with a baby is a simple and natural task, researchers have worked long and hard to describe the distinct acoustic properties of IDS. This article shows that a great deal of progress has been made. Some acoustic properties of IDS, including aspects of vocal pitch and rhythm, are now well-established. Other acoustic properties that pertain to the vocal resonances of the speech signal, are less well understood and are currently a focus of intense research attention and debate. Explaining exactly how specific properties of IDS impact infant development is another challenge that continues to drive research activity. In this article, we also outline some of the knowledge gaps that are energizing researchers to reach for a deeper understanding of the unique acoustic properties of IDS and to explore how IDS is connected with infant speech. As we learn more about IDS and why babies thrive on it, we are also finding ways to leverage this knowledge to promote infant development.

Infants Prefer Infant-Directed Speech

IDS has captivated scientists precisely because it is so effective in enticing infant attention. Across many studies in

Figure 1. *Publications (**blue**) and citations (**orange**) of papers on infant-directed speech (IDS) from 1990 to 2020. From the Web of Science.*



https://doi.org/10.1121/AT.2021.17.1.26

which infants are presented a choice to listen to samples of IDS and adult-directed speech (ADS), infants (even newborns) repeatedly show a clear and strong preference to listen to IDS, with few studies deviating from this pattern. A meta-analysis found that the average listening time difference between IDS and ADS (or "the effect size of IDS preference" in statistical terms) was significant and large (Dunst et al., 2012).

Infant preference for IDS, being recognized as one of the most robust behaviors measured in infancy, was selected as the target behavior in a large-scale study designed to understand how subject variables and testing methodologies affect the measurement of infant behavior. This study, conducted by the ManyBabies Consortium, involved 67 laboratories across North America, Europe, Australia, and Asia. The findings provided further conclusive evidence of infants' preference for IDS over ADS (ManyBabies Consortium, 2020). There is no doubt that infants are attracted to IDS.

Acoustic Properties of Infant-Directed Speech

What is it about IDS that babies like? Studies show that when caregivers talk to their infant, they modify their speech on multiple levels. This includes basic speech patterns that play a broad role in communication and can be observed across different languages (conveying emotion and talker information and basic units such as vowels, consonants, and word forms) as well as acoustic cues that mark specific lexical, grammatical, and pragmatic features that are important in a specific language. Our focus here is on basic acoustic speech patterns that have a broad impact and are more likely to be universal across languages.

To understand the acoustic properties of IDS, it is useful to know that the acoustic speech signal has two independent components, referred to as the source and the filter. The vocal source component is determined by how fast the vocal folds vibrate, which determines the voice pitch or fundamental frequency (see article on singing by Sundberg et al. on pages 43-51). The voice pitch of an infant or child is much higher than that of an adult because their short, light vocal folds vibrate faster compared with the longer and thicker vocal folds of an adult. Talkers also vary their voice pitch by adjusting the tension of the vocal folds.

The vocal filter component refers to the effects of the length and shape of the vocal tract, the term used to refer to the tube formed by the vocal folds on one end and the mouth at the other end. Movements of the tongue, jaw, and lips vary the length and shape of the vocal tract, which determines the resonances of the vocal tract.

The acoustic patterns formed by the vocal resonances created when we speak are referred to as formants and are numbered in ascending frequency value (the lowest is the 1st formant [F1], next is the 2nd formant [F2], etc.). The formants are essentially narrow frequency regions where acoustic energy is increased because these frequencies vibrate most easily within the associated vocal tract space. The first three formants contain critical acoustic information for speech communication.

The vocal resonances and associated formant frequencies are higher for the short vocal tract of an infant or child compared with the longer vocal tract of an adult. Talkers modify the resonance of the vocal tract to create different vowel sounds by moving their articulators to create different vocal tract shapes, such as by adjusting the degree and location of constrictions along the vocal tract.

An extensive body of research has concentrated on describing the acoustic structure of IDS. This work has considered each component, typically by comparing samples of IDS with comparable samples of ADS (Soderstrom, 2007).

Voice Pitch and Rhythmic Properties of Infant-Directed Speech

The distinct vocal source properties of IDS are well-established (see **Multimedia** 1-5 at acousticstoday.org/polkamedia for audio examples in English and in Turkish; see video example at bit.ly/3m3ecHh). Overall, higher voice pitch, wider voice pitch range and greater pitch variability have been found in IDS (compared with ADS) in a variety of languages, including both nontonal languages (Fernald et al., 1989) and tonal languages (Liu et al., 2009). Several studies have shown that high voice pitch is the primary acoustic determinant of the infants' preference for IDS (Fernald and Kuhl, 1987; Leibold and Werner, 2007). Research focused on the speech movements that occur during IDS have observed, as expected, that adults produce faster vocal fold vibrations and also raise their larynx when they talk to young infants (Ogle and Maidment, 1993; Kalashnikova et al., 2017). Larynx raising naturally occurs when vocal fold tension increases (which raises voice pitch) and can also shorten the overall vocal tract length.

It is widely held that the primary goal or intention guiding these characteristic voice pitch properties is conveying emotion to the young infant (Saint-Georges et al., 2013). Understanding the emotional expression in IDS led researchers to explore the pitch contours found in IDS. Fernald and Simon (1984) observed that most utterances in IDS had either rising or falling pitch contours. Stern and colleagues (1982) identified the social and linguistic context where these pitch contours were used. For example, a rising contour was frequently used when mothers tried to engage in eye contact with an inattentive baby. Studies also show that creating "happy talk" is the fundamental goal of IDS and that positive affect is what drives infant preference (Singh et al., 2002). Thus, understanding pitch contours in IDS can help us decode the affective function of IDS.

In terms of rhythmic features, IDS universally contains shorter utterances and longer pauses between words; in some languages, including English and Japanese, there is also an enhanced lengthening of words or syllables at the end of a phrase or utterance (Fernald et al., 1989; Martin et al., 2016). This is helpful because natural fluent speech typically lacks pauses between words, something you notice when encountering an entirely foreign language. This also highlights an initial challenge for babies, learning which speech patterns are reoccurring words, aka word segmentation. Infants begin to acquire word segmentation skills at around 6 months, through experience listening to a specific language and before they attach meaning to each word they hear (Jusczyk, 1999).

Overall, the tempo of IDS provides the infant with a speech stream that is easier to track with clearer cues marking word boundaries and other syntactic units. Consistent with this, the most prominent rhythm in the acoustic speech signal, which matches the timing of stressed syllables, was observed to be stronger in IDS compared with ADS (Leong et al., 2017). This speech rhythm was also prominent (and synchronized) in mother and infant brain patterns when they watched a nursery rhyme video together (Santamaria et al., 2020).

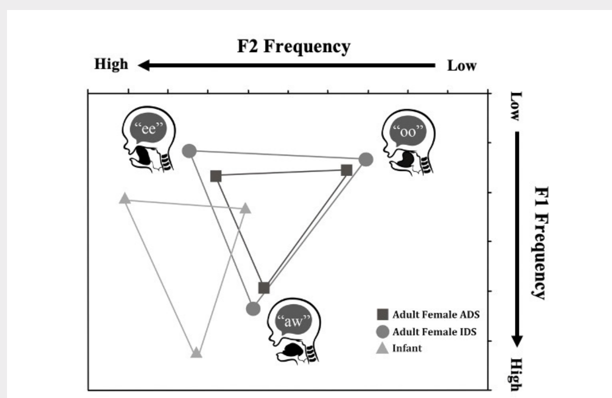
The enhanced temporal properties of IDS likely explains the positive effects of IDS on infant speech processing. For example, IDS facilitates infant word segmentation performance (Thiessen et al., 2005), an important language-specific speech-processing skill that emerges in infancy and supports early word learning. The list of the positive effects of IDS rhythm on speech processing, which includes supporting better discrimination and tracking of syllable patterns and detection of speech in noise, continues to grow (Soderstrom, 2007).

Vocal Resonance Properties of Infant-Directed Speech

Research on IDS has also considered the other fundamental component of speech, the filter or resonance properties. The focus here has been on vowel sounds. Early research by Kuhl and colleagues (1997) reported that vowels are produced in an exaggerated form in IDS; this *hyperarticulation* of vowels expands the vowel space, a standard graphic display that captures how vowel articulation and formant patterns are related.

In the classic vowel space (**Figure 2**), F1 increases as the tongue/jaw height decreases and F2 increases as the tongue constriction moves to the front of the mouth. Importantly, three vowel sounds found in every spoken language, "ee," "aw," and "oo," form the corners of this F1/F2 vowel space. These corner vowels are associated with gestural extremes that define the full range of movements that we use to create vowel sounds: "ee" has the most high and front constriction of the vocal tract, "oo" has the most high and back constriction of the vocal tract, and "aw" has the most open and unconstricted posture of the vocal tract. All other vowel sounds fall within the limits defined by these corner vowel sounds.

Figure 2. *The articulatory/acoustic vowel space corresponding to vowels produced by an adult female in ADS (**squares**) and in IDS (**circles**), and by an infant (**triangles**). F1 and F2, 1st and 2nd formants, respectively.*



The finding that the vowel space is larger, with the corner vowels spaced further apart in IDS compared with ADS, is also illustrated schematically in **Figure 2**.

Increasing the acoustic distance between vowels enhances recognition of distinct vowel sounds. Moreover, even in ADS, a larger vowel space is typically associated with more intelligible speech (Bradlow et al., 1996). An expanded vowel space was observed for IDS vowels produced in several languages, suggesting this is a universal feature of IDS (Kuhl et al., 1997). Moreover, infants with a mother who expanded her vowel space when producing IDS also performed better in a speech sound discrimination task (Liu et al., 2009). Other work suggests that early exposure to vowel expansion in IDS is associated with better expressive and receptive language skills at two years (Hartman et al., 2016).

The idea that caregivers expand their vowel space to make speech clearer is consistent with the finding that adults do not expand their vowel space when speaking to a pet with little or no capacity for acquiring language skills even though pet-directed speech typically contains the characteristic pitch properties of IDS that convey affect as outlined above (Burnham et al., 2002). IDS appears to be a form of hyperarticulated speech that promotes language development by clarifying and enhancing speech segments (e.g., vowels and consonants).

We now see this view as incomplete. As work advanced, an expanded vowel space in IDS has not been found in all languages, in all interactions, or at all infant ages (see Hartman et al., 2016). Vowel expansion has also been absent when studies relied on samples of natural spontaneous speech instead of the structured laboratory-recorded samples used in earlier studies (Martin et al., 2015). In a study of IDS in Japanese, Miyazawa and colleagues (2017) observed vowel space expansion when average formant values were considered. However, the IDS vowel sounds were actually not more distinct because there was much more acoustic variability and overlap among different vowels produced in IDS, which makes recognizing distinct vowel sounds more, not less, difficult.

Actually, some researchers suggested that vowel space expansion in IDS is an unintended side effect of the increased pitch (which also raises the larynx and shortens the vocal tract) and slower speaking rate that characterizes this speech style rather than being shaped by the parents' direct effort to clarify speech patterns (McMurray et al., 2013). For example, in her study of IDS in Dutch, Benders (2013) observed a reduced rather than an expanded vowel space, and she also observed an overall rise in all formant frequencies for IDS vowels relative to ADS vowels. This pattern is noteworthy given that smiling, which shortens the vocal tract length, is known to shift formants upward and further apart in frequency (Shor, 1978). Thus, Benders claimed that caregivers modify their vowel sounds in IDS, especially with young infants, to communicate positive emotion rather than to provide clearer speech. Meanwhile, Miyazawa and colleagues (2017) observed that IDS had a breathier voice quality compared with ADS vowels, an effect also associated with communicating emotion.

Yet another perspective emerged from a study by Kalashnikova and colleagues (2017) that directly examined the tongue and lip movements that caregivers make when they produce IDS and ADS, which are the source of these formant patterns. In this study, special sensors were strategically placed on eight moms to track their lip and tongue movements while they produced IDS with their 11-month-old infants and ADS with the experimenter. Surprisingly, when the moms produced the corner vowels in IDS, their tongue and lips movements were not exaggerated or hyperarticulated as the researchers expected, even though simultaneous acoustic recordings showed expansion of the IDS vowel space.

Through an analysis that combined lip movement and vowel formant measures, Kalashnikova et al. (2017) inferred that the IDS speech was produced with a shorter vocal tract and higher pitch, which mothers can create by raising their larynx. They concluded that expansion of the F1/F2 acoustic vowel space in IDS is not created by the mother's intentional efforts to produce clearer vowel sounds using exaggerated articulatory movements. Instead, mothers are speaking with the unintentional purpose of sounding smaller and thus unthreatening and nonaggressive.

This speaking style can also be viewed as the mother trying to imitate her infant. Moreover, this form of *vocal social convergence* is observed in other species and is regarded as a mechanism for creating a close emotional bond between adult and infant, presumably to ensure that infant offspring survive and thrive. Kalashnikova et al. (2017) claim that in early infancy, any benefits of IDS related to clarifying speech units are secondary to this basic social/emotional bonding goal, and, as in evolution, these linguistically motivated patterns likely emerge later in development and piggy-back on to this social bonding function.

Overall, what is happening to the resonance component of the speech signal when caregivers use IDS is not fully resolved. Caregivers may be modifying their speech to clarify speech units and boost language development, to convey positive emotions, to sound smaller and build social bonds, or some combination of these effects. Although the details remain unclear, understanding how these modifications impact infant development continues to ignite and steer ongoing research.

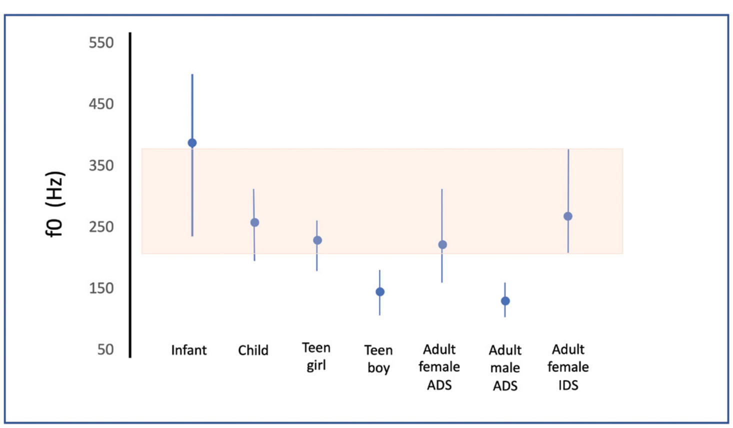
Infant-Directed Speech and Infant Speech: An Important Connection?

As outlined in **Vocal Resonance Properties of Infant-Directed Speech**, there are different viewpoints regarding what motivates the use of IDS. One idea to emerge recently is that when mothers use IDS, they are altering their speech to sound smaller and more like an infant. Although this is a new perspective on IDS, the act of unconsciously adapting your speech to mirror or imitate features of your conversational partner is not a new observation. This has been noted and studied extensively in adult speech communication and is often referred to as *phonetic convergence*. Moreover, in adult-to-adult interaction, speech convergence is typically associated with liking or holding a positive attitude toward your conversational partner (Pardo, 2013).

Other findings point to an important connection between IDS and infant speech. First, there are indeed clear parallels between IDS and infant speech. With respect to vocal source properties, infant speech and IDS have similar voice pitch values, particularly when IDS is produced by a female adult/mother. **Figure 3** shows voice pitch values across the life span, including voice pitch values for IDS produced by female adults. **Figure 3**, *pink box*, highlights the range in which voice pitch values overlap across infant speech and speech produced by an adult female using IDS.

Although voice pitch values can overlap across IDS and infant speech, the vocal filter properties of infant speech and IDS are more distinct. When an adult female raises her larynx and spreads her lips to shorten her vocal tract length, she will sound like a smaller person. Nevertheless, a mother cannot shorten her vocal tract enough to match the vocal tract length of her infant. Infant speech has much higher vocal resonances, reflected in the formant frequencies uniquely associated with a talker with a very short vocal tract. This results in higher formant frequency values for infant speech compared with adult speech which are shown by a spectrogram of the vowel "ee" produced by an infant and a female adult (**Figure 4**). These differences are also apparent in the vowel space shown in **Figure 2**, where you can see that the corner

Figure 3. *Typical average voice pitch (f0) values for speakers across the life span. Blue lines, observed range of values observed within each group. Pink box, voice pitch range where infant and adult female IDS values overlap. Data from Masapollo et al., 2016, Table 1.*



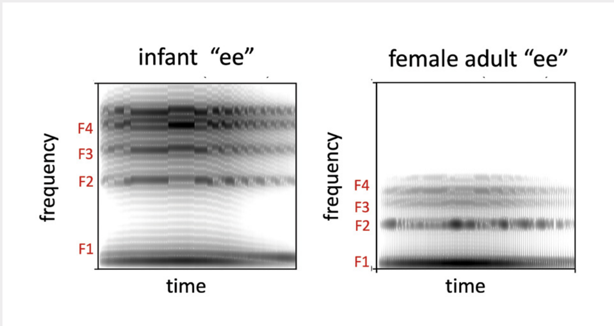


Figure 4. *Example spectrograms of "ee" vowels showing the formant patterns produced by an infant (**left**) and an adult female (**right**).*

vowels found in infant speech are acoustically distinct from those found in adult female ADS or IDS vowels. That being said, when mothers use IDS, they do their best to alter each component of speech to approximate or converge with the properties of an infant talker.

Notably, these changes in IDS align very well with what infants like when it comes to speech. Infants not only favor IDS, they also are attracted to infant speech. In listening preference tests, young infants listened longer to vowels sounds produced by an infant over vowel sounds produced by a female adult (Masapollo et al., 2016). To measure this, researchers created vowel sounds that simulate a 6-month-old talker using a special speech synthesizer (examples are shown in **Figure 4**). This study also showed that infants prefer each component of infant speech; infants prefer the high voice pitch of an infant and also the high-frequency vocal resonances produced by a small-infant vocal tract. Importantly, this means that infants have a distinct preference for infant talkers; they are not simply favoring a high voice pitch that is known to be an attractive property of IDS.

It was also noted in this study (Masapollo et al., 2016) that some infants vocalized and smiled more when they listened to infant vowel sounds compared with adult vowel sounds, suggesting that the strong attraction to infant speech may stimulate and reward vocal exploration in young infants. Another study (Polka et al., 2014) examined the infants' ability to recognize the same vowel when it is produced by different talkers (e.g., man, woman, child, and infant). Including infant vowel sounds in this task made it more challenging, but it also boosted infant listening times and recognition performance (Polka et al., 2014). It seems that infant speech sounds grab and hold infant attention in ways that help babies recognize important speech categories. Overall, these findings suggest that vocal convergence in IDS may play a broader role beyond social bonding. Vocal convergence may also help the infant discover that their own vocalizations are part of this vocal social space and motivate them to explore and refine their vocal skills.

Infants' attraction to infant speech sounds raises new questions. Are mothers instinctively aware of this bias? Does this motivate them to sound more infant-like when using IDS? Then again, maybe mothers are shaping this bias by using vocal convergence? Is there an important connection between IDS and infant speech that will help us understand how infants acquire spoken language? These intriguing questions drive current research and promise to shed new light on the role of IDS in infant development.

Multitasking with Infant-Directed Speech

Although we have learned a great deal about the acoustic properties of IDS and how it affects infant speech processing, we are just beginning to understand how IDS impacts infant development. As noted in **Vocal Pitch and Rhythmic Properties of Infant-Directed Speech**, **Vocal Resonance Properties of Infant-Directed Speech**, and **Infant-Directed Speech and Infant Speech: An Important Connection?**, a range of functions for IDS has been proposed, including attracting and holding infant attention, highlighting and enhancing linguistic segments and structure, communicating emotion, strengthening infant/caregiver social bonds, and stimulating vocal exploration (Saint-Georges et al., 2013; Golinkoff et al., 2015).

It is widely agreed that IDS is a powerful multitasking tool that caregivers flexibly adapt to meet the moment-to-moment needs of their infant. This adaptability is ideal for meeting parent and infant needs but presents challenges for scientific investigation. In specific contexts, the diverse functions shaping IDS are often intertwined in complex ways (Saint-Georges et al., 2013). For example, modifications that communicate positive affects can promote social bonding while also facilitating speech processing by enhancing attention.

Moreover, these different functions are not equally prominent in any given interaction or across all ages or developmental stages. For example, in IDS with young infants (<12 months), communicating emotion is often more prominent than clarifying linguistic structures. In IDS with older children (>12 months), the reverse occurs, such that highlighting linguistic structure is often more prominent than communicating emotion and building social bonds. No doubt, IDS is best understood in the context of infant/caregiver interaction and when the needs of the child and the intentions of the caregiver are identified.

Contingency and Synchrony Are Fundamental

IDS is recognized to be dynamic and actively shaped by both the infant and caregiver. Contingent and synchronized responding between mother and infant is a core feature of IDS. Although an IDS speaking style can be simulated by an adult, IDS production is facilitated by the presence of a baby. The salience of caregiver responsiveness is demonstrated by the finding that adults can readily identify audio recordings of IDS recorded with and without an infant present (Trehub et al., 1997).

Saint-Georges and colleagues (2013) proposed that IDS creates an interactive communication loop connecting the infant and the caregiver in a synergistic way. This idea has motivated researchers to search for physiological markers of enhanced synchrony during IDS. Synchronous activity has been observed in heart rate and respiration measures (McFarland et al., 2019) and gaze patterns (Santamaria et al., 2020) recorded during parent/infant interactions where IDS is commonly used.

The powerful role of dynamic social interaction is also reinforced by research showing that infants can readily learn to discriminate consonants from a foreign language in a live interaction involving IDS but not from audiovisual recordings (Kuhl et al., 2003). It is also intriguing to consider how the musical quality of IDS (which is enhanced in infant-directed singing) shapes this parent-infant synchrony, given that early music exposure affects infant brain development (Zhao and Kuhl, 2020).

The critical role of IDS contingency and synchrony is also supported by evidence that challenges on each side of the interactional loop affect the synergistic connection created via IDS. For example, from the caregiver side, mothers with depression tend to include less affective information and have smaller pitch variations when speaking to their infants (Kaplan et al., 2001). Infants' learning is affected when maternal depression persists over an extended period (Kaplan et al., 2011). However, infants of depressed mothers remain responsive to IDS from nondepressed fathers and the quality of IDS is soon improved when the mother's depression is lifted (Kaplan et al., 2004). On the infant side, the preference for IDS is absent or reduced among children with autism spectrum disorder, presumably reflecting difficulties in processing the heightened emotional content of IDS (Kuhl et al., 2005).

New Directions

Going forward, research is moving quickly to expand our knowledge of IDS. Although we have learned a great deal about the acoustic properties of IDS, we need to learn more about the speech movements that give rise to IDS signals. This type of work is technically challenging but critical for understanding exactly what caregivers are doing when they adapt their speech for their infant, especially with respect to vocal resonance properties.

Future research will also continue to build a more complete understanding of the social, emotional, cognitive, and linguistic benefits of IDS for the developing child. Research exploring the physiological responses of interacting caregivers and infants will play a central role by helping us identify and understand the contingent and synchronous processes that are mediated by IDS. Each new finding pushes our curiosity to a higher level. We are confident that IDS will hold the interest of infants, caregivers, and scientists for a long time and can help us understand conditions that compromise parent/infant connection and identify new ways to optimize infant development.

Acknowledgments

We thank Claire Ying Ying Liu for preparing Figure 2 and Sandra Trehub and the ManyBabies Consortium for sharing infant-directed speech samples.

References

Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development* 36, 847-862. https://doi.org/10.1016/j.infbeh.2013.09.001.

Bradlow, A. R., Torretta, G. M., and Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication* 20, 255-272. https://doi.org/10.1016/S0167-6393(96)00063-5.

Burnham, D., Kitamura, C. M., and Vollmer-Conna, U. (2002). What's new pussycat? On talking to babies and animals. *Science* 296, 1435. https://doi.org/10.1126/science.1069587.

Dunst, C., Gorman, E., and Hamby, D. (2012). Preference for infant-directed speech in preverbal young children. *Center for Early Literacy Learning* 5(1), 1-13. Available at http://www.earlyliteracylearning.org/cellreviews/cellreviews_v5_n1.pdf.
Fernald, A., and Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development* 10(3), 279-293. https://doi.org/10.1016/0163-6383(87)90017-8.
Fernald, A., and Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology* 20(1), 104-113. https://doi.org/10.1037/0012-1649.20.1.104.
Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., and Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language* 16(3), 477-501. https://doi.org/10.1017/S0305000900010679.
Golinkoff, R. M., Can, D. D., Soderstrom, M., and Hirsh-Pasek, K. (2015). (Baby) talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science* 24(5), 339-344. https://doi.org/10.1177/0963721415595345.
Hartman, K. M., Bernstein-Ratner, N., and Newman, R. S. (2016). Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language* 44 (5), 1-26. https://doi.org/10.1017/S0305000916000520.
Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Science* 3(9), 323-328. https://doi.org/10.1016/S1364-6613(99)01363-7.
Kalashnikova, M., Carignan, C., and Burnham, D. (2017). The origins of babytalk: smiling, teaching or social convergence? *Royal Society Open Science* 4(8), 170306. https://doi.org/10.1098/rsos.170306.
Kaplan, P. S., Bachorowski, J. A., Smoski, M. J., and Zinser, M. (2001). Role of clinical diagnosis and medication use in effects of maternal depression on infant-directed speech. *Infancy* 2(4), 537-548. https://doi.org/10.1207/S15327078IN0204_08.
Kaplan, P. S., Danko, C. M., Diaz, A., and Kalinka, C. J. (2011). An associative learning deficit In 1-year-old infants of depressed mothers: Role of depression duration. *Infant Behavior and Development* 34(1), 35-44. https://doi.org/10.1016/j.infbeh.2010.07.014.
Kaplan, P. S., Dungan, J. K., and Zinser, M. C. (2004). Infants of chronically depressed mothers learn in response to male, but not female, infant-directed speech. *Developmental Psychology* 40, 140-148. http://doi.org/10.1037/0012-1649.40.2.140.
Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., and Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science* 277(5326), 684-686. https://doi.org/10.1126/science.277.5326.684.
Kuhl, P. K., Coffey-Corina, S., Padden, D., and Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Developmental Science* 8(1), F1-F12. https://doi.org/10.1111/j.1467-7687.2004.00384.x.
Kuhl, P. K., Tsao, F. M., and Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America* 100(15), 9096-9101. https://doi.org/10.1073/pnas.1532872100.
Leibold, L. J., and Werner, L. A. (2007). Infant auditory sensitivity to pure tones and frequency-modulated tones. *Infancy* 12(2), 225-233. https://doi.org/10.1111/j.1532-7078.2007.tb00241.x.
Leong, V., Kalashnikova, M., Burnham, D., and Goswami, U. (2017). The temporal modulation structure of infant-directed speech. *Open Mind* 1(2), 78-90. https://doi.org/10.1162/OPMI_a_00008.
Liu, H. M., Tsao, F. M., and Kuhl, P. K. (2009). Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and five-year-old children: A longitudinal study. *Journal of Child Language* 36(4), 909-922. https://doi.org/10.1017/S030500090800929X.
ManyBabies Consortium (2020). Quantifying sources of variability in infancy research using the infant-directed speech preference. *Advances in Methods and Practices in Psychological Science* 3(1), 24-52, https://doi.org/10.1177/2515245919900809.
Martin, A., Igarashi, Y., Jincho, N., and Mazuka, R. (2016). Utterances in infant-directed speech are shorter, not slower. *Cognition* 156, 52-59. https://doi.org/10.1016/j.cognition.2016.07.015.
Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., and Cristia, A, (2015). Mothers speak less clearly to infants than to adults: A comprehensive test of the hyperarticulation hypothesis. *Psychological Science* 26(3), 341-347. https://doi.org/10.1177/0956797614562453.
Masapollo, M., Polka, L., and Menard, L. (2016). When infants talk, infants listen: Pre-babbling infants prefer listening to speech with infant vocal properties. *Developmental Science* 19(2), 318-328. https://doi.org/10.1111/desc.12298.
McFarland, D. H., Fortin, A. J., and Polka, L. (2019). Physiological measures of mother–infant interactional synchrony. *Developmental Psychobiology* 62(1), 50-61. https://doi.org/10.1002/dev.21913.
McMurray, R., Kovack-Lesh, K. A., Goodwin, D., and McEchron, W. (2013). Infant-directed speech and the development of speech perception: Enhancing development or an un-intended consequence? *Cognition* 129(2), 362-378. https://doi.org/10.1016/j.cognition.2013.07.015.
Miyazawa, K., Shinya, T., Martin, A., Kikuchi, H., and Mazuka, R. (2017). Vowels in infant-directed speech: More breathy and more variable, but not clearer. *Cognition* 166, 84-93. https://doi.org/10.1016/j.cognition.2017.05.003.
Ogle, S. A., and Maidment, J. A. (1993). Laryngographic analysis of child-directed speech. *International Journal of Language & Communication Disorders* 28(3), 289-297.
Pardo, J. (2013). Measuring phonetic convergence in speech production. *Frontiers in Psychology* 4, 559. https://doi.org/10.3389/fpsyg.2013.00559.
Polka, L., Masapollo, M., and Menard, L. (2014). Who's talking now? Infants' perception of vowels with infant vocal properties. *Psychological Science* 25(7), 1448-1456. https://doi.org/10.1177/0956797614533571.
Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., Laznik, M. C., and Cohen, D. (2013). Motherese in interaction: at the cross-road of emotion and cognition? (A systematic review). *PLoS ONE* 8(10), e78103. https://doi.org/10.1371/journal.pone.0078103.
Santamaria, L., Noreika, V., Georgieva, S., Clackson, K., Wass, S., and Leong, V. (2020). Emotional valence modulates the topology of the parent-infant inter-brain network. *Neuroimage* 207, 116341, 1-17 https://doi.org/10.1016/j.neuroimage.2019.116341.
Shor, R. (1978). The production and judgement of smile magnitude. *The Journal of General Psychology* 98(1), 79-96. https://doi.org/10.1080/00221309.1978.9920859.
Singh, L., Morgan, J. L., and Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy* 3(3), 365-394. https://doi.org/10.1207/S15327078IN0303_5.
Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review* 27(4), 501-532. https://doi.org/10.1016/j.dr.2007.06.002.
Stern, D. N., Spieker, S., and MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology* 18(5), 727-735. https://doi.org/10.1037/0012-1649.18.5.727.

Thiessen, E. D., Hill, E. A., and Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy* 7(1), 53-71. https://doi.org/10.1207/s15327078in0701_5.

Trehub, S .E., Unyk, A. M., Kamenetsky, S. B., Hill, D. S., Trainor, L. J., Henderson, J. L., and Saraza, M. (1997). Mothers' and fathers' singing to infants. *Developmental Psychology* 33, 500-507. https://doi.org/10.1037/0012-1649.33.3.500.

Zhao, C. T., and Kuhl, P. K. (2020). How early music training changes the brain. *Acoustics Today* 16(3). 61-69. https://doi.org/10.1121/AT.2020.16.3.61.

About the Authors

Linda Polka
linda.polka@mcgill.ca

School of Communication Sciences and Disorders
Centre for Research on Brain, Language, and Music
Faculty of Medicine and Health Sciences
McGill University
2001 McGill College Avenue
Montréal, Quebec H3A 1G1, Canada

Linda Polka is professor and graduate program director in the School of Communication Sciences and Disorders, McGill University (Montréal, QC, Canada) and director of the McGill Infant Speech Perception Lab. She is also a fellow of the Acoustical Society of America and was recently (2019) chair of the Technical Committee for Speech Communication. Her work explores how infant speech perception is shaped by universal biases and language experience in both monolingual and bilingual infants.

Yufang Ruan
yufang.ruan@mail.mcgill.ca

School of Communication Sciences and Disorders
Centre for Research on Brain, Language, and Music
Faculty of Medicine and Health Science
McGill University
2001 McGill College Avenue, Room 845
Montréal, Quebec H3A 1G1, Canada

Yufang Ruan is a doctoral student in the Infant Speech Perception Lab, McGill University (Montréal, QC, Canada). Her research interests concern language development and developmental disorders. She received her MSc from Beijing Normal University (China) and her BA from Dalian University of Technology (China).

Psychoacoustics of Tinnitus: Lost in Translation

Christopher Spankovich, Sarah Faucette, Celia Escabi, and Edward Lobarinas

Tinnitus: What Is It?

Tinnitus is the perception of sound without an external source, often experienced as a constant or frequent ringing, humming, or buzzing. Tinnitus is reported by more than 50 million people in the United States alone (Shargorodsky et al., 2010); conservatively 1 in 10 US adults has tinnitus (Bhatt et al., 2016). It is estimated that 20-25% of patients with tinnitus consider the symptoms to be a significant problem (Seidman and Jacobson, 1996).

Various nomenclature has been applied to describe tinnitus, including terms such as subjective or objective tinnitus and the more recent recommended terms primary and secondary tinnitus (Tunkel et al., 2014). Primary tinnitus refers to tinnitus that is idiopathic and may or may not be associated with sensorineural hearing loss (SNHL; hearing loss (HL) related to dysfunction of the inner ear and auditory nerve). Secondary tinnitus refers to tinnitus that is associated with a specific underlying cause other than SNHL or an identifiable organic condition such as pulsatile tinnitus (heartbeat perception in ear). Our discussion here is focused on primary tinnitus, which is the more common variant.

Causes of Tinnitus

The understanding of the physiological source of primary tinnitus has significantly expanded over the past 30 years. Numerous portions of the auditory pathway and nonauditory neural pathways have been implicated in tinnitus perception and reaction. Still, the exact mechanisms underlying tinnitus remain elusive.

Contemporary research points to both peripheral and central factors that underlie tinnitus. In other words, peripheral changes to the auditory part of the inner ear and auditory neural integrity, most commonly resulting from noise exposure, ototoxic drugs, and age-related factors, result in compensatory changes/neural plasticity at more central segments of the pathway. These changes include (1) an increase spontaneous neural activity of excitatory neurons/neurotransmitters and a reciprocal decrease in activity of inhibitory neurons/neurotransmitters, resulting in central gain; (2) distortions in frequency representation as input to more central regions is restricted due to peripheral damage; and (3) nonauditory pathway/structure recruitment, suggesting a multisensory and distributed brain network implicated in mediating tinnitus perception and reaction. Simply stated, tinnitus is the attempt of the brain to fill in the reduced peripheral input (Spankovich, 2019).

Perception Versus Reaction to Tinnitus

A critical distinction is the perception of tinnitus versus the reaction to tinnitus. The tinnitus percept or phantom sound itself has minimal repercussions for morbidity or mortality. Conversely, the reaction or emotional response to tinnitus can have a substantial effect on a person's functional status (Jastreboff and Hazell, 1993). Almost everyone with tinnitus, whether bothersome or not, would want the percept eliminated if possible (Tyler, 2012).

Clearly, tinnitus is not perceived as a positive experience. The onset of tinnitus perception does not generally evoke a feeling of improved health or well-being. For example, if you hear a grinding noise in your car engine one day, your first reaction is not positive in nature. The reaction to tinnitus may further be influenced by events related to its onset, where the tinnitus becomes a reminder of that experience (Fagelson, 2007). For example, a person with an acoustic neuroma (tumor of the auditory-vestibular nerve) and its associated tinnitus may experience an enhanced awareness to tinnitus changes and an exacerbated reaction due to concern that it is a sign the tumor is growing larger or more invasive. A soldier who has experienced tinnitus during or following an active engagement may be

https://doi.org/10.1121/AT.2021.17.1.35

reminded of that experience by the presence of the tinnitus, reinforcing in their mind that they cannot escape the tinnitus nor escape or leave the past behind them.

Is There a Cure for Tinnitus?

Despite decades of research there is no “cure” for tinnitus. Indeed, no medication or surgery can remove the tinnitus perception from the brain. In the absence of a cure, medical interventions focus on mitigating the tinnitus reaction. Treatment options generally include some form of counseling (e.g., education on the neuroscience of tinnitus) and use of sound enrichment (e.g., hearing aids) to help diminish the tinnitus perception and reaction (Tunkel et al., 2014).

The most common side effects of tinnitus are sleep disturbances, concentration issues, loss of quiet/feeling of inability to escape the tinnitus, and emotional/stress-based issues (Tyler and Baker, 1983). Although rare, tinnitus can result in suicidal ideation and suicide (Szibor et al., 2019). It is also common for persons with tinnitus to attribute their hearing difficulties to their tinnitus perception (Henry et al., 2015). This is in general unsupported; tinnitus does not cause HL, but, rather, HL causes tinnitus. Nonetheless, tinnitus can affect concentration that can impact listening (Burns-O'Connell et al., 2019) and speech understanding with competing noise (Oosterloo et al., 2020).

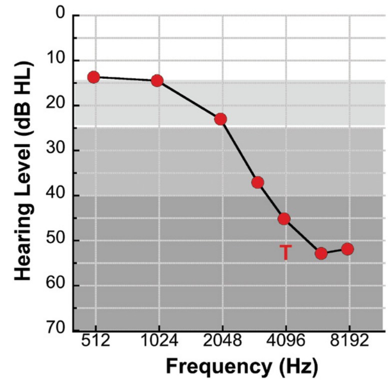


Figure 1. *Audiogram and tinnitus match. An audiogram is a graphical representation of hearing thresholds at different frequencies. The frequency of the signal is plotted against the level required for the patient to detect the sound stimulus.* ***Gray bars****, variable levels of hearing loss (HL). Levels less than 15 dB HL are considered normal hearing and greater than 15 dB HL corresponds to escalating levels of HL severity. This tinnitus patient (right ear only) has normal low-frequency hearing, sloping to a high-frequency HL or a moderate level. Red T represents the pitch and loudness match for the patient. In brief, this patient matched their tinnitus pitch at 4,000 Hz and loudness at 55 dB HL or 10 dB sensation level to their threshold (45 dB HL).*

Measuring Tinnitus in Humans

There is currently no widely accepted or validated method to identify the presence of primary tinnitus and quantification of its perceptual characteristics other than what is reported by the patient. An objective measure of primary tinnitus by the clinician, a long-held goal, is complicated by the relationship among tinnitus, HL, and hyperacusis (sound sensitivity related to increased central neural activity compensating for reduced peripheral input) and a lack of sensitivity and specificity from electrophysiological measures or imaging studies. Developing objective measures of tinnitus has been challenging in studies in both human and animal.

Then again, perhaps an objective measure to rule-in or rule-out the presence of tinnitus is not necessary. For example, the gold standard for assessment of hearing sensitivity is the pure-tone audiogram (**Figure 1**), which indicates the lowest sound level a human or animal can detect at different frequencies. The audiogram is, however, a psychophysical measure that is nonobjective in nature. Of course, a method to measure tinnitus that has translation between animal models and humans would be most efficacious to empower development of diagnostic and treatment approaches.

Recommendations for the psychophysical assessment of tinnitus were postulated over 20 years ago by the Ciba Foundation and the National Academy of Sciences. Methods for administering the psychoacoustic battery for tinnitus assessment have been reviewed (Henry, 2016). To date, standardization of these procedures is still not recognized; however, generalized clinical methods are briefly described here.

Pitch match (PM) measures the patient’s perceived tinnitus pitch (perception of sound frequency) by matching the tinnitus to a specific frequency or range of frequencies. PM is typically measured, although crudely, using an audiometer where two sounds are played, and the person with tinnitus chooses the pitch closer to their tinnitus percept. Most patients with peripheral hearing deficits match their tinnitus pitch at frequencies respective to their HL.

Loudness match (LM) measures the perceived loudness level of the tinnitus and is typically reported in sensation level (SL), that is, the level relative to the individual's auditory threshold. LM is measured by presenting a tone or noise and asking the patient to indicate if the tone is softer or louder than their tinnitus. The intensity of the tone is then adjusted until the patient reports the tone is a comparable loudness to their tinnitus.

To determine the ability of external sounds as a means of tinnitus suppression (i.e., masking), the minimum masking level (MML) is often assessed. The MML is the minimum level of external acoustic stimulation, typically a noise, needed to cover up the patient's tinnitus perception. In this assessment, a low-level broadband noise (BBN) is presented to the patient via an audiometer. The intensity of the noise is slowly raised until the patient reports they can no longer hear their tinnitus. This measure can be useful in prescribing sound therapy-based recommendations to patients.

Last, many patients report that tinnitus suppression may persist after the masker has been turned off, a phenomenon known as residual inhibition (RI). RI is a measure of the duration of patient-reported tinnitus suppression after a patient has been presented with masking noise. Noise is presented at 10 dB above MML for 60 s. After the procedure, the patient is asked if they experience any difference in their tinnitus. If their tinnitus is suppressed, the duration of suppression is timed; this often only lasts for seconds to minutes. Use of this procedure is cautioned, however, because it can exacerbate tinnitus in some patients.

The use of these subjective tinnitus measurements is not ubiquitous among audiologists The reliability of such measurements is often questioned, as is their purpose, and none of these measurements are recommended in the American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS) (Tunkel et al., 2014) nor the US Department of Veterans Affairs (Henry and Manning, 2019) clinical guidelines. At best, the tinnitus assessment serves to provide a quantification of a person's tinnitus perception, which, in turn, may provide a tool in counseling and considerations of sound-based therapy. Nonetheless, the tinnitus assessment does not necessarily differentiate persons with tinnitus and those feigning a tinnitus perception. Furthermore, the results of the tinnitus assessment described have limited correlation with tinnitus reaction (Manning et al., 2019).

Tinnitus or No Tinnitus?

In 2006, Jim Henry and colleagues at the National Center for Rehabilitative Auditory Research (NCRAR) in Portland, OR, described an automated system for the psychoacoustic assessment of tinnitus. The system was a self-assessment tool using on-screen instructions that allowed the individual with tinnitus to alter frequency and intensity parameters to match the psychoacoustic attributes of their tinnitus percept.

In an interesting twist, the study design included a group of participants with comparable hearing status but that did not report tinnitus, rather they were instructed to feign a tinnitus percept. Henry et al. (2006) reported significant differences in outcomes between the tinnitus and the no-tinnitus group. First, the loudness matches for the tinnitus group were two to four times greater than the no-tinnitus group. Second, the PM was nearly an octave higher for the tinnitus group relative to the no-tinnitus group. Reliability between sessions was not different for the LM, but the no-tinnitus group showed greater variance for PM. The authors proposed developing a statistical method to determine the probability an individual has tinnitus based on variance of the measures.

Perception Versus Reaction

An additional limitation of the psychophysical assessment of tinnitus is the lack of correlation to tinnitus reaction or functional impact (Manning et al., 2019). Numerous scales exist to measure tinnitus reaction (Meikle et al., 2012). Although the relationship between measures of tinnitus perception and reaction is weak, this does not eliminate their relative potential for determining the presence of tinnitus and identifying the affective and functional impact, respectively. Finally, visual numeric rating scales (NRS) and visual analog scales (VAS) to assess tinnitus loudness are additional methods to quantify tinnitus perception. However, studies suggest that rather than correlating to loudness, these measures are more reflective of the tinnitus reaction (Hall et al., 2017).

Measuring Tinnitus in Animals

Animal models of tinnitus are important for more invasive measures to determine physiological changes related to tinnitus perception and development of potential therapeutics. Early animal studies (Jastreboff et al., 1988) used high doses of sodium salicylate, the active ingredient in aspirin, to induce transient tinnitus. Aspirin at high doses has been shown to reliably induce tinnitus in humans but is also usually reversible and again limited to high doses; a baby aspirin is unlikely to cause tinnitus.

Given that tinnitus is a phantom auditory perception, how can it be measured in animals? The simple answer is that patients cannot perceive quiet while tinnitus is present, and neither can animals. Across studies, animals are trained to exhibit one set of behaviors (e.g., pressing levers, moving from one side of the chamber to another, climbing a pole) when there is no sound in the environment and another set of behaviors when sound is on in order to obtain food or avoid punishment. Among the animal models (Brozoski and Bauer, 2016), the most common approach is to have animals (usually rodents) detect a gap in a continuous sound. When tinnitus is present, animals make more errors detecting gaps in continuous sound, especially if the frequency of the continuous sound is similar in pitch to their tinnitus.

Several of these animal studies have shown that the pattern of results supports the presence of tinnitus after high doses of sodium salicylate, quinine (an antimalarial drug known to induce tinnitus in humans), and noise exposure. Importantly, the pitch of the tinnitus is consistent with the adjusted frequency range (relative to peripheral HL) reported in humans.

To effectively test animals for the presence of tinnitus, several fundamental features are necessary for rigorous investigation. These include the use of well-established behavioral response paradigms for determining the phantom sound of tinnitus, known and reliable inducers of tinnitus, and/or reliable physiological responses consistent with the presence of tinnitus. Psychophysical assessment of tinnitus is typically categorized either as an interrogative model, which evaluates changes in behavioral outcomes as a function of tinnitus, or as a reflexive model, which assess changes in automatic, lower-order processing responses consistent with the perception of a phantom sound.

Interrogative models require that the animal voluntarily respond to the acoustic environment indicating the presence of silence or the presence of an auditory stimulus. Early preclinical behavioral measures of tinnitus used interrogative methods, operant conditioning, and response suppression to detect and characterize the presence of tinnitus (Jastreboff et al., 1988). In the first animal model, rats were conditioned to associate a mild but unavoidable foot shock that occurred after a continuous sound was turned off. This resulted in suppressed licking from a water spout in preparation of the imminent shock. Following conditioning, rats in the experimental group were given a high dose of sodium salicylate, whereas the control group received a placebo. During this phase, the foot shock was eliminated but the sound conditions remained. Rats in the control group continued to suppress licking when the sound was turned off because the lack of sound was associated with foot shock. In contrast, rats treated with sodium salicylate continued to lick even when the sound was turned off. Simply put, the animals could not tell that the sound was turned off (presumably due to presence of tinnitus) and continued to lick from the waterspout.

A number of subsequent animal models have shown results consistent with the presence of tinnitus and consistent with Jastreboff's lick suppression model (Eggermont and Roberts 2015). Other models have used either avoidable shock or positive reinforcement with food whereby animals have to differentiate between trials with sound and trials with no sound. Although interrogative assessments in animal models are crucial for investigating perceptual correlates of tinnitus, it is important to note the considerable challenges in interrogative models because behavioral conditioning requires lengthy and consistent training schedules (Brozoski and Bauer, 2016), and even then, some animals may not respond as expected due to inability to do the task or lack of motivation.

Given the challenges associated with interrogative models, reflexive models for tinnitus assessment have been widely used for determining the presence of tinnitus. The acoustic startle reflex (ASR) is a large-motor response akin to a jumping/jolt-like response that can be readily elicited in rodents using a loud startling acoustic stimulus. The ASR can be easily measured in rodents using pressure sensitive platforms to record the amplitude and duration of the reflex (Turner et al., 2006).

Interestingly, the ASR can be attenuated by presenting an acoustic cue before the startling acoustic stimulus. For example, a 50-ms tone before the loud startling stimulus will result in a reduction in the ASR. Because of the compressed time frame, the changes in the ASR are believed to involve rapid lower level auditory processing before the startle elicitor; in other words, the animal did not need to think it over before startling. For the purposes of assessing tinnitus, a continuous sound is played in the background and a brief gap is presented before the loud startling stimulus, called gap prepule inhibition of an acoustic startle (GPIAS). However, if tinnitus is present and the background continuous sound is similar in pitch to the tinnitus, the animals will be unable to reliably detect the gap and there will be no reduction or smaller reductions in the ASR. This paradigm can be used to assess both the presence of tinnitus as well as the frequency range of the tinnitus. For example, Lobarinas et al. (2015) demonstrated that rats with evidence of noise-related tinnitus based on the ASR showed an improved startle response (i.e., less tinnitus filling gap) when treated with the drug cyclobenzaprine (a tricyclic antidepressant).

Reflexive models such as the GPIAS have the main advantage of precluding overt and long behavioral training. However, these models are not without their limitations, such as habituation of the ASR (Lobarinas et al., 2013a) and loss of reactivity to loud startling stimuli following unilateral HL. Although these drawbacks have called the widespread use of the GPIAS into question, it remains the most popular paradigm used in preclinical models of tinnitus. One way to overcome one of the limitations of the GPIAS is to elicit the startling response with a tactile stimulus. Thus, an acoustic stimulus can be used to cue the imminent startling stimulus without concerns of the efficacy of an acoustic startle elicitor. Lobarinas et al. (2013a) demonstrated success using an air puff to the animal's back to elicit a robust startle response. Cuing the air puff with an acoustic stimulus reduced the startle response to the air puff. Using a tactile stimulus such as the air puff has allowed the model to be used to study unilateral and bilateral tinnitus as well as other auditory phenomena such as hearing in noise and suprathreshold deficits associated with subclinical HL (Lobarinas et al., 2017).

Lost in Translation

Animal and human findings relative to tinnitus often have conflicting results. For example, the idea that tinnitus fills in perception of a silent gap works in animals but is not so clear that it does so in humans. Continued improvements in animal models will make it possible to evaluate physiological correlates and basic mechanisms under controlled tinnitus-inducing conditions as well as to evaluate hypotheses generated from studying human participants.

It is also worth noting that these animal models of tinnitus all focus on the perception; no animal models of the affective/emotional reaction to tinnitus are well accepted. Here we will consider two tinnitus-related phenomena that have been lost in translation between animals and humans: (1) tinnitus filling in a silent gap and (2) how peripheral hearing damage creates tinnitus.

Tinnitus "Filling in the Gap"

The application of gap detection and suppression of a startle reflex has become a common high-throughput model of tinnitus assessment in animals. In simple terms, the paradigm suggests that the presence of tinnitus disrupts the ability of the animal to detect the silent gap, and thus the startle response is less suppressed. Attempts to translate this measure to humans has been less promising.

For example, Fournier and Hebert (2013) used a GPIAS model measuring reflexive eye blink activity in participants with tinnitus compared with controls. They observed that participants with tinnitus had decreased inhibition of eye blink activity when it was preceded by a silent gap in noise compared with control participants. Nonetheless, despite all tinnitus participants reporting high-pitch ringing tinnitus, the decreased inhibition was found for both low- and high-frequency noise stimuli. In other words, the decreased inhibition was not limited to gaps in noise reflective of the tinnitus perception (high frequency). The findings contradicted the assertion that tinnitus is simply filling in the gap and frequency-specific deficits observed in some animal models but did show altered ASR of eye blinking.

In the same year, Campolo et al. (2013) performed a similar study of tinnitus filling in the gap but alternatively focused on perception of a silent gap than on an effect on an ASR. Applying methods comparable to animal experiments (50-ms silent intervals in varying noise bands), they observed no deficits in detecting the silent gap in persons with or without tinnitus. Similar findings were reported by Boyen et al. (2015), including no difference in detecting shorter gap durations.

The difference in findings of these studies may be explained by different neural circuits underlying reflexive responses and behavioral-/perception-based responses. Fournier and Hebert (2013) were relying on a startle reflex (eye blink) compared with a conscious perception of a sound (or silent gap) as in Campolo et al. (2013) and Boyen et al. (2015).

Loss of Tuning

One of the earliest proposed theories of tinnitus initiation was the discordant damage theory. According to this theory (an extension of theories proposed by Tonndorf 1981a,b), the outer hair cells (OHCs) of the mammalian cochlea are more prone to damage than the inner hair cells (IHCs), resulting in imbalanced activity via type I and type II afferent fibers that, respectively, carry signals from the ear to the dorsal cochlear nucleus (DCN), the first auditory center in the brain. The alteration of input to the DCN results in loss of inhibition and compensatory mechanisms at more central sites, including bursting neural activity, mapping reorganization, decreased inhibition, and central gain mentioned in **Tinnitus: What Is It?**.

Kaltenbach and Afman (2000) showed that significant IHC damage can prevent the onset of hyperactivity in the DCN. Tonndorf's (1981) original model suggested a decoupling of stereocilia (the hair-like projections from the cell) between the OHCs and the tectorial membrane (a membrane floating above the hair cells) that leads to loss of energy and increased noise at the level of the hair cell underlying tinnitus generation. Tonndorf's follow-up theory (1987) suggested that tinnitus was equivalent to chronic pain in the somatosensory system and a result of preferential damage to the OHCs and established an analogy of tinnitus to chronic pain.

In contrast to the discordant damage theory, cochlear insults that commonly lead to chronic tinnitus in humans have been found to produce a long-term decrease in the auditory neuronal spontaneous activity (Liberman and Dodds, 1984). Tinnitus is strongly correlated with HL and cochlear damage as a result of ototoxicity or noise exposure. Specifically, IHC/synaptic loss has been speculated to produce tinnitus.

To explore this relationship, a behavioral gap detection task was used to determine the presence of tinnitus in a chinchilla model with selective IHC loss following administration of carboplatin. Carboplatin is an ototoxic anticancer drug known to cause significant IHC loss (>80% loss) while leaving OHCs largely intact (<5% loss) in the chinchilla, an effect unique to the chinchilla model (Lobarinas et al., 2013b). Preliminary data showed overall poorer gap detection performance when tested at lower presentation levels, but the findings were not frequency specific. The absence of frequency-specific deficits suggested that these animals did not perceive tinnitus even with severe IHC loss. Thus, IHC damage alone does not seem sufficient to generate tinnitus and support the discordant dysfunction theory of tinnitus or a combination of OHC and IHC/synapse injury at play.

Changes to psychophysical tuning curves may offer insight into differentiating OHC vs. IHC/synaptic contributions to the onset of tinnitus but are currently limited to humans in regard to tinnitus effects. A psychophysical tuning curve is a method that can be used to generate comparable data to the physiological frequency threshold curve for a single auditory nerve fiber. A narrowband noise of variable center frequency is used as a masker, and a fixed frequency and fixed-level pure tone at about 20 dB HL is commonly the target. The level of masker is found that just masks the tone for different masker frequencies. With OHC damage, the tuning curve becomes flattened and less sharp due to loss of sensitivity.

For example, Tan et al. (2013) examined psychophysical tuning curves in persons with HL and tinnitus and in persons with HL and no tinnitus. Both groups were compared with a reference group of persons with normal hearing. The normal-hearing group showed expected patterns of low thresholds and sharp tuning curves; these patterns are thought to reflect the nonlinearity of the OHCs. Interestingly, the HL group with tinnitus showed better thresholds, greater residual compression, and better tuning than the no-tinnitus group in the midfrequency range. This was likely reflective of the greater high-frequency HL of the tinnitus group relative to the no-tinnitus group that had a wider array of patterns. Thus, the finding could simply reflect differences in hearing thresholds; however, after matching participants based on HL, the pattern persisted. Tan et al. suggested that the findings may be explained by the tinnitus group having residual OHC function and a preferential loss of IHCs or afferents.

The difference in the animal model of widespread loss of IHCs and lack of tinnitus evidence compared with psychoacoustic tuning curves in humans implicating IHCs/synapse may also be explained by the discordant damage theory. The carboplatin model creates a pure loss of IHCs/synapses without damage to OHCs. Still, humans may still have some level of damage to their OHCs not reflected in their tuning curves. In other words, it would be parsimonious to suggest that there is likely a ratio of damage to both hair cell types involved and necessary to generate tinnitus perception. Currently, animal-based versions of psychophysical tuning curves are lacking. Development of this paradigm in preclinical models would provide an opportunity to further advance tinnitus research and enhance translation.

Challenge to Psychoacousticians

Psychophysical measures of tinnitus are numerous. In general, these measures have been applied to match attributes of tinnitus, determine the affective impact of tinnitus, and identify the site of lesion and subtyping of tinnitus physiological origin. It is apparent that most psychoacoustic measures such as PM, LM, and MML do not reliably correlate with measures of tinnitus reaction. The use of numerical rating scales, visual analog scales, and questionnaires on affective elements appear to best capture elements of tinnitus reaction.

Tools to assess affective elements have been established in humans but represent a challenge for animal models. The important question is, do animals experience tinnitus related distress? From clinical data, the majority of individuals who experience tinnitus are not disabled by it. It is thus reasonable to expect only a minority of animals will be debilitated by tinnitus. To address this issue, a large number of animals would be needed in studies of tinnitus-related distress, with careful consideration of confounding variables (e.g., housing, animal handling).

The overarching question, given the state of the science, is how can we use principles from psychophysics to identify one or several measures of tinnitus using perceptual attributes of tinnitus that can differentiate individuals who actually experience tinnitus from persons with reported tinnitus but no actual tinnitus perception? Furthermore, how can we use psychophysical experiments to better inform our understanding of the tinnitus neurophysiology. With improved models, further progress can be attained to lead to novel therapeutics for the management of tinnitus.

References

Bhatt, J. M., Lin, H. W., and Bhattacharyya, N. (2016). Prevalence, Severity, Exposures, and Treatment Patterns of Tinnitus in the United States. *Journal of the American Medical Association Otolaryngology-Head & Neck Surgery* 142(10), 959-965.

Boyen, K., Baskent, D., and van Dijk, P. (2015). The gap detection test: Can it be used to diagnose tinnitus, *Ear and Hearing* 36(4), 138-145.

Brozoski, T. J., and Bauer, C. A. (2016). Animal models of tinnitus. *Hearing Research* 338, 88-97.

Burns-O'Connell, G., Stockdale, D., and Hoare, D. J. (2019). Soldiering on: a survey on the lived experience of tinnitus in aged military veterans in the UK. *Medical Humanitarian* 45(4), 408-415.

Campolo, J., Lobarinas, E., and Salvi, R. (2013). Does tinnitus "fill in" the silent gaps? *Noise and Health* 15(67), 398-405.

Eggermont, J. J., and Roberts, L. E. (2015). Tinnitus: Animal models and findings in humans. *Cell Tissue Research* 361(1), 311-336.

Fagelson, M. A. (2007). The association between tinnitus and posttraumatic stress disorder. *American Journal of Audiology* 16(2), 107-117.

Fournier, P., and Hebert, S. (2013). Gap detection deficits in humans with tinnitus as assessed with the acoustic startle paradigm: Does tinnitus fill in the gap? *Hearing Research* 295, 16-23.

Hall, D. A., Mehta, R. L., and Fackrell, K. (2017). How to choose between measures of tinnitus loudness for clinical research? A report on the reliability and validity of an investigator-administered test and a patientreported measure using baseline data collected in a phase IIa drug trial. *American Journal of Audiology* 26(3), 338-346.

Henry, J. A. (2016). "Measurement" *of* tinnitus. *Otology and Neurotology* 37(8), e276-e285.

Henry, J. A., and Manning, C. (2019). Clinical protocol to promote standardization of basic tinnitus services by audiologists. *American Journal of Audiology* 28(1S), 152-161.

Henry, J. A., Griest, S., Zaugg, T. L., Thielman, E., Kaelin, C., Galvez, G., and Carlson, K.F. (2015). Tinnitus and hearing survey: a screening tool to differentiate bothersome tinnitus from hearing difficulties. *American Journal of Audiology* 24(1), 66-77.

Henry, J. A., Rheinsburg, B., Owens, K. K., and Ellingson, R. M. (2006). New instrumentation for automated tinnitus psychoacoustic assessment. *Acta Oto-Laryngologica* 126(Suppl. 556), 34-38.

Jastreboff, P. J., and Hazell, J. W. (1993). A neurophysiological approach to tinnitus: clinical implications. *British Journal of Audiology* 27(1), 7-17.

Jastreboff, P. J., Brennan, J. F., and Sasaki, C. T. (1988). An animal model for tinnitus. *Laryngoscope* 98(3), 280-286.

Kaltenbach, J. A., and Afman, C. E. (2000). Hyperactivity in the dorsal cochlear nucleus after intense sound exposure and its resemblance to tone-evoked activity: A physiological model for tinnitus. *Hearing Research* 140(1-2), 165-172.

Liberman, M. C., and Dodds, L. W. (1984). Single-neuron labeling and chronic cochlear pathology. II. Stereocilia damage and alterations of spontaneous discharge rates. *Hearing Research* 16(1), 43-53.

Lobarinas, E., Blair, C., Spankovich, C., and Le Prell, C. (2015). Partial to complete suppression of unilateral noise-induced tinnitus in rats after cyclobenzaprine treatment. *Journal of the Association of Research in Otolaryngology* 16(2), 263-272.

Lobarinas, E., Hayes, S. H., and Allman, B. L. (2013a). The gap-startle paradigm for tinnitus screening in animal models: Limitations and optimization. *Hearing Research* 295, 150-160.

Lobarinas, E., Salvi, R., and Ding, D. (2013b). Insensitivity of the audiogram to carboplatin induced inner hair cell loss in chinchillas. *Hearing Research* 302, 113-120.

Lobarinas, E., Spankovich, C., and Le Prell, C. G. (2017). Evidence of "hidden hearing loss" following noise exposures that produce robust TTS and ABR wave-I amplitude reductions. *Hearing Research* 349, 155-163.

Manning, C., Thielman, E. J., Grush, L., and Henry, J. A. (2019). Perception versus reaction: Comparison of tinnitus psychoacoustic measures and tinnitus functional index scores. *American Journal of Audiology* 28(1S), 174-180.

Meikle, M. B., Henry, J. A., Griest, S. E., Stewart, B. J., Abrams, H. B., McArdle, R., Myers, P. J., Newman, C. W., Sandridge, S., Turk, D. C., and Folmer, R. L. (2012). The tinnitus functional index: Development of a new clinical measure for chronic, intrusive tinnitus. *Ear and Hearing* 33(2), 153-176.

Oosterloo, B. C., Homans, N. C., and Goedegebure, A. (2020). Tinnitus affects speech in noise comprehension in individuals with hearing loss. *Otology & Neurotology* 41(9), e1074-e1081.

Seidman, M. D., and Jacobson, G. P. (1996). Update on tinnitus. *Otolaryngology Clinics of North America* 29(3), 455-465.

Shargorodsky, J., Curhan, G. C., and Farwell, W. R. (2010). Prevalence and characteristics of tinnitus among US adults. *American Journal of Medicine* 123(8), 711-718.

Spankovich, C. (2019). *Tinnitus and Sound Sensitivity* 3rd ed. In J. A. Galster (Ed.),: Audiology Treatment Thieme. New York, NY,, pp. 328-355.

Szibor, A., Makitie, A., and Aarnisalo, A. A. (2019). Tinnitus and suicide: An unresolved relation. *Audiology Research* 9(1), 222.

Tan, C. M., Lecluyse, W., McFerran, D., and Meddis, R. (2013). Tinnitus and patterns of hearing loss. *Journal of the Association of Research in Otolaryngology* 14(2), 275-282.

Tonndorf, J. (1981a). Stereociliary dysfunction, a case of sensory hearing loss, recruitment, poor speech discrimination and tinnitus. *Acta Oto-Laryngologica* 91(5-6), 469-479.

Tonndorf, J. (1981b). Tinnitus and physiological correlates of the cochleo-vestibular system: Peripheral; central. *Journal of Laryngology Otology Supplement* 4, 18-20.

Tonndorf, J. (1987). The analogy between tinnitus and pain: A suggestion for a physiological basis of chronic tinnitus. *Hearing Research* 28(2-3), 271-275.

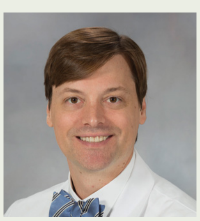
Tunkel, D. E., Bauer, C. A., Sun, G. H., Rosenfeld, R. M., Chandrasekhar, S. S., Cunningham, E. R., Jr., Archer, S. M., Blakley, B. W., Carter, J. M., Granieri, E. C., and Henry, J. A. (2014). Clinical practice guideline: Tinnitus. *Otolaryngology-Head and Neck Surgery* 151(Suppl. 2), S1-S40.

Turner, J. G., Brozoski, T. J., Bauer, C. A., Parrish, J. L., Myers, K., Hughes, L. F., and Caspary, D. M. (2006). Gap detection deficits in rats with tinnitus: A potential novel screening tool. *Behavioral Neuroscience* 120(1), 188-195.

Tyler, R. S. (2012). Patient preferences and willingness to pay for tinnitus treatments. *Journal of the American Academy of Audiology* 23(2), 115-125.

Tyler, R. S., and Baker, L. J. (1983). Difficulties experienced by tinnitus sufferers. *Journal of Speech and Hearing Disorders* 48(2), 150-154.

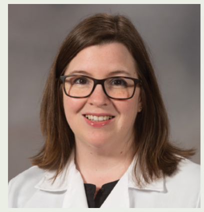
About the Authors



Christopher Spankovich
cspankovich@umc.edu

Department of Otolaryngology—Head and Neck Surgery
University of Mississippi Medical Center
Jackson, Mississippi 39216, USA

Christopher Spankovich is an associate professor and vice chair of research in the Department of Otolaryngology—Head and Neck Surgery, University of Mississippi Medical Center (Jackson). He obtained his MPH from Emory University (Atlanta, GA), AuD from Rush University (Chicago, IL), and PhD from Vanderbilt University (Nashville, TN). He is a clinician-scientist with a translational research program focused on the prevention of acquired forms of hearing loss, tinnitus, and sound sensitivity. He continues to practice clinically, with a special interest in tinnitus, sound sensitivity, ototoxicity, hearing conservation, and advanced diagnostics. He serves as an associate editor for the *International Journal of Audiology*.



Sarah Faucette
sfaucette@umc.edu

Department of Otolaryngology—Head and Neck Surgery
University of Mississippi Medical Center
Jackson, Mississippi 39216, USA

Sarah Faucette earned both her PhD in communication sciences and disorders and her AuD from East Carolina University (Greenville, NC). She is an assistant professor of audiology in the Department of Otolaryngology—Head and Neck Surgery, University of Mississippi Medical Center (Jackson). She also has an appointment in the Memory Impairment and Neurodegenerative Dementia (MIND) Center, where she is the lead research audiologist for the Aging and Cognitive Health Evaluation in Elders (ACHIEVE) trial. Her research interests include amplification, tinnitus, and hearing loss in the aging population.

Celia D. Escabi
celia.escabi@utdallas.edu

School of Behavioral and Brain Sciences
University of Texas at Dallas
Dallas, Texas 75235, USA

Celia D. Escabi earned both her bachelor's degree and AuD from the University of South Florida (Tampa). She is currently pursuing a PhD in hearing science at the University of Texas at Dallas. Celia is a certified and licensed audiologist for the state of Texas and works as a research audiologist under her research mentors Colleen Le Prell and Edward Lobarinas. Her current research goals include advancing her basic science skills and bridging them together with her clinical experiences. Celia's primary research interests include auditory cognitive neuroscience and pharmacological interventions for acquired auditory disorders.

Edward Lobarinas
edward.lobarinas@utdallas.edu

School of Behavioral and Brain Sciences
University of Texas at Dallas
Dallas, Texas 75235, USA

Edward Lobarinas earned a bachelor's degree from Rutgers University (New Brunswick, NJ) and a master's degree and PhD from the State University of New York at Buffalo. He has had faculty appointments in audiology programs at the University at Buffalo, the University of Florida (Gainesville), and the University of Texas at Dallas. He is trained as a clinical audiologist and a basic researcher. His research interests include tinnitus, tinnitus treatments, perceptual changes associated with the selective loss of inner hair cells, and machine learning applications for assistive listening devices. His work has been funded by the National Institutes of Health, private foundations, and industry.

One Singer, Two Voices

Johan Sundberg, Björn Lindblom, and Anna-Maria Hefele

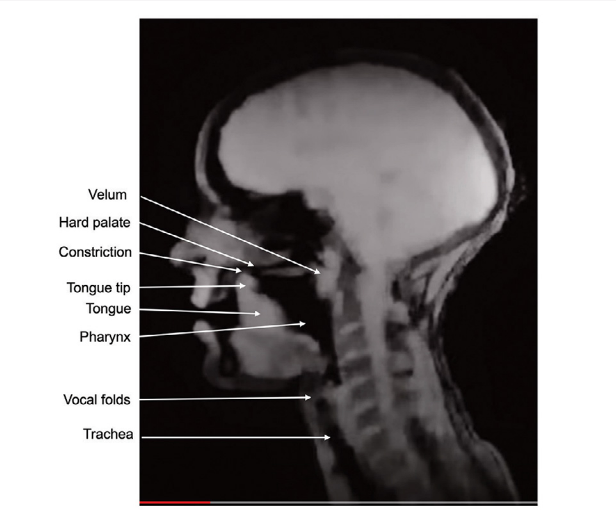
Introduction: Rendering Melodies with Overtones

A single singer but two voices? Experience that situation by visiting world-voice-day.org/EDU/Movies and check the second movie with the title "*Sehnsucht nach dem Frühlinge* (Mozart) — Anna-Maria Hefele (AMH). There, coauthor AMH sings a song by Mozart, first with her singing voice and then with two simultaneous voices, a drone (a low-pitched, continuously sounding tone) plus a whistle-like high-pitched tone that renders the melody. How is this possible? That is the question that we pose here. Let us start by recalling how sounds are created by the instrument AMH is playing, the human voice.

Vocal Sound Generation

Figure 1 shows a frame from the movie mentioned above. It shows a magnetic resonance imaging (MRI) with the various parts of the voice organ labeled. Voice production is the summed result of three processes: (1) compression of air below the vocal folds; (2) vocal fold vibration, quasi-periodically chopping airflow from the subglottal region; and (3) filtering of the acoustic signal of this pulsatile airflow.

Figure 1. *Magnetic resonance (MR) image of Anna-Maria Hefele's (AMH's) head and throat, taken from the video where she performs a Mozart melody in overtone singing technique.*



The overpressure of air below the folds throws them apart, thus allowing air to pass through the slit between them. Then, aerodynamic conditions reduce the air pressure along the folds, which, together with the elasticity of their tissue, closes the slit. The same pattern is then repeated, thus generating vocal fold vibration.

The vibration generates a pulsatile airflow as seen in **Figure 2A**, producing sound, the voice source. The pitch is determined by the vibration frequency, whereas the waveform is far from sinusoidal. Hence, this airflow signal is composed of a number of harmonic partials. In other words, the frequency of a partial number (n) = $n \times f_o$, where f_o is the frequency of the lowest partial, the fundamental or vibration frequency. The amplitudes of the partials tend to decrease with their frequency; the amplitude of n tends to be something like 12 dB stronger than the amplitude of $n \times 2$. The spectrum envelope of the voice source is rather smooth and has a negative slope as seen in **Figure 2B**.

The voice source is injected into the vocal tract (VT), which is a resonator. Hence it possesses resonances at certain frequencies. Partials with frequencies close to a VT resonance frequency are enhanced and partials further away are attenuated (see **Figure 2C**). Therefore, the spectrum envelope of the sound radiated from the lip opening (**Figure 2A**) contains peaks at the VT resonance frequencies and valleys in-between them. In this sense, the VT resonances form the spectrum envelope of the sound emitted to the free air. Probably for this reason, VT resonances are frequently referred to as formants.

The frequencies of the formants are determined by the shape of the resonator composed of the pharynx and the mouth cavities, the VT. For example, the VT length has

https://doi.org/10.1121/AT.2021.17.1.43

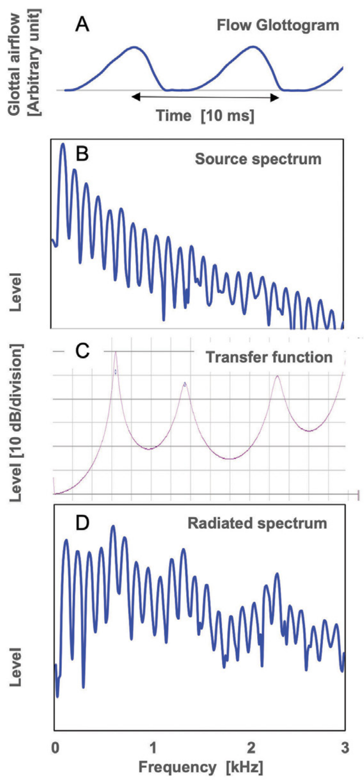


Figure 2. ***A*** *and* ***B****: typical waveform and spectrum, respectively, of the glottal airflow during phonation.* ***C*** *and* ***D****: vocal tract transfer function for the vowel /ae/ and the corresponding radiated spectrum, respectively.*

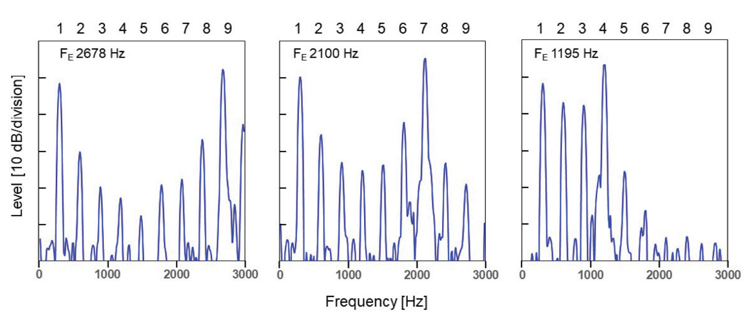
a strong effect on these frequencies and protruding the lips makes the VT longer, thus lowering the formant frequencies. The shape of the VT can be varied within wide limits. Moreover, by bulging the tongue more or less and in various directions, the VT can be narrowed or widened in almost any place along its length axis from the deep pharynx to the hard palate. Also, the jaw and lip openings contribute to determining the VT shape. As a result, the formant frequencies can be varied within quite wide ranges: the first formant between about 150 and 1,000 Hz, the second from about 500 and 3,000 Hz, and the third from about 1,500 Hz and 4,000 Hz.

Overtone Singing

What is overtone singing, then? The term covers several different styles. Overtone singing was described as early as the nineteenth century by a famous singing teacher Manuel Garcia (Wendler, 1991) and has attracted the interest of several researchers. Smith and colleagues (1967) described "throat singing," a type of chant performed by Tibetan Lamas, (see, e.g., s.si.edu/37QCSOZ). It is produced by males with special types of vocal fold vibrations, referred to as vocal fry register. Its pitch is stable and very low and is produced by a vocal fold vibration pattern in which every second or third airflow pulse is attenuated. Consequently, the pitch period of this drone is doubled or tripled. A similar type of phonation often occurs in phrase endings in conversational speech but is then typically aperiodic. In throat singing, two of the overtones are quite strong, and audible, thereby together giving the impression of a "chord." Throat singing is regarded as sacred in some Asian cultures.

The overtone singing demonstrated by AMH in the above link can be produced by both females and males. However, the fundamental frequency of the drone is not as low as in throat singing. In AMH's case, it is in the range typical of female speaking voices. The melody is played in a much higher pitch range by very strong overtones. **Figure 3** shows some examples where overtones number 9, 7, and 4 are the strongest in the spectrum.

Figure 3. *Examples of spectra produced in overtone singing by AMH. F_E, frequency of the enhanced overtone.*



In everyday listening, overtones are not perceived individually. Instead, the patterning of their amplitudes, determined by the resonance characteristics of the VT, collectively contributes to what our auditory system perceives as the "timbre" or vocal "color" of what is sung or spoken. The reason why overtones normally escape us is linked to the way our hearing works. It processes spectral contents by averaging the information in broad frequency bands, the so-called critical bands (Moore, 2012).

Characteristics of Overtone Singing

Against this background, we may be forgiven for finding overtone singing a rather puzzling phenomenon. Here we present an attempt to shed some light on its phonatory, articulatory, and acoustic bases.

We begin with a sample from AMH's rendering of the Mozart melody. **Figure 4A** presents the first few bars of the beginning of Mozart's theme in musical notation. Let us take a moment to consider what it would take to sing this sequence using overtones?

Figure 4 gives an answer in-principle. The format of the musical score is used to indicate the timing and pitch of each overtone. Along the frequency scale, the first eight harmonics are drawn at equidistant intervals. Hypothetically, let us suppose that the singer selects a fundamental frequency near 300 Hz for the drone. That implies a "keyboard" of the following frequencies of the first eight overtones: #2, 600 Hz; #3, 900 Hz; #4,1,200 Hz; #5,1,500; #6, 1,800 Hz; #7, 2,100 Hz; and #8, 2,400 Hz.

Note the following ratios: (1) 1,500/1,200 = 1.25; (2) 1,800/1,200 = 1.5; (3) 2,400/1,200 = 2. Relative to the first note at 1,200 Hz, the ratios correspond to a major third, a perfect fifth, and an octave, respectively. Those intervals are the ones needed to produce the notes of the first two bars.

For readers who find that result just a little too convenient, we should point out that it is no accident. Our musical scales and their intervals bear a very close evolutionary relationship to the physical structure of periodic sounds. Such sounds are constituted by the partials, the frequencies of which form a harmonic series (Gill and Purves, 2009). This implies that the musical intervals octave, fifth, fourth, major third, and minor third appear between the six lowest spectrum partials. Hence, it is possible to play melodies with the partials of a constant drone tone.

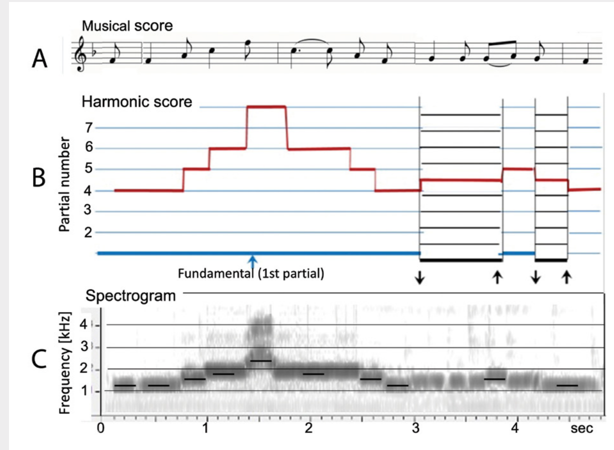


Figure 4. *Schematic illustration of singer AMH's overtone singing performance of a Mozart song.* ***A:*** *the musical score.* ***B:*** *partials used for the melody tones (**red lines**) and for the drone (**blue** and **black lines**).* ***C:*** *wideband spectrogram of the performance.*

However, in the third bar in **Figure 4A** with chord tones C E G Bb, the pitch G appears. The series just mentioned does not provide an equally convenient choice for that note. Looking ahead to AMH's performance (**Figure 4B**), we see how she handles the situation: She lowers the fundamental of the drone so as to produce an overtone whose relationship to 1,200 Hz is that of a major second or about 1,125 Hz! Our harmonic score includes that approach as illustrated by a shift in the fundamental and harmonics of the drone.

Figure 4C shows real data, a spectrogram of one of AMH's overtone singing versions of the song. To enhance the display of the overtones, we show a wideband filtering that portrays the overtones as dark patches. To help interpret their positions, we added black short lines that indicate the expected frequency values on the assumption that the 4th, 5th, 6th, and 8th harmonics of a 300-Hz fundamental serve as the melodic building blocks.

We note that these predictions parallel AMH's overtones rather well. However, the black marks slightly underestimate the observed values. Why? AMH used a fundamental frequency slightly lower than our hypothetical 300 Hz.

This example illustrates the fact that overtone singing derives from the lawful way in which the harmonics are organized in periodic sounds. Overtone singers are able to exploit this patterning. They have developed a way of selecting and amplifying the harmonics and have refined their VT motor skills to be able, with great precision in time and frequency, to produce harmonics as melodic sequences. Next, we test the hypotheses that (1) enhancing and selecting a single partial is the result of VT shapes producing clustering of formants; and (2) that overtone singing is produced with a regular sound source.

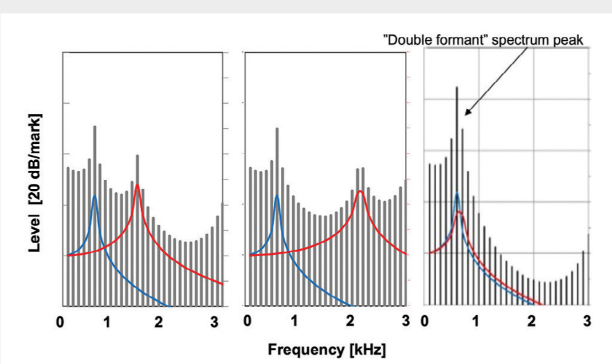
Acoustic Theory

Enhancing Single Harmonics

How can formants produce the excessive amplitudes of the single overtones illustrated in **Figure 3**? Mathematically, the spectral envelope, the function determining the amplitudes of the overtones, is the sum of formant resonance curves (which vary with changes in VT shape) and certain factors such as glottal source and radiation characteristics (which do not depend on articulatory activity). The only input to the calculation is the frequencies and bandwidths of the formants. The latter factor generally varies with the former factor in a predictable manner so formant amplitudes need not be specified. **Figure 5** illustrates this predictability.

Figure 5 shows three line spectra with the cardinal shapes of the resonance curves for the first and second formants (henceforth F1 and F2). The amplitudes of the partials and their spectral envelopes were derived in accordance with the standard source-filter model (Fant, 1960). This theory treats the envelope as the sum of formant curves and the constant contributions of source and radiation characteristics, which are not shown in **Figure 5**. The bandwidths are secondary aspects, being determined mainly by the frequencies of the formants (Fant, 1972).

Figure 5. *Schematic illustration of the spectrum effects of moving the frequencies of two formants closer together.* ***Vertical lines****, partials of a drone with a fundamental frequency of 100 Hz;* ***blue*** *and* ***red curves****, first (F1) and second (F2) formants, respectively.* ***Left:*** *F1 = 600 Hz, F2 = 1,400 Hz.* ***Center:*** *F1 = 600 Hz; F2 = 2,150 Hz.* ***Right:*** *F1 = 600 Hz; F2 = 650 Hz, thus creating a "double formant," that creates a very strong partial (****arrow****).*



In **Figure 5**, F1 was fixed at 600 Hz while F2 was varied. When F1 and F2 approach almost identical frequencies (**Figure 5**, *right*), creating, as it were, a *double formant*, their individual peaks merge into a single maximum, with a significant increase in the amplitude of the closest partial. In other words, acoustic theory states that formant amplitudes are predictable and thus suggests an answer to the question asked in the first sentence of this section. Enhancing the amplitude of individual overtones is possible: *Move two formants close to each other in frequency. Create a double formant!*

Measurements and Modeling

Vocal Tract Shapes

As shown above, if the formant frequencies are determined by the shape of the VT, so what was the shape of AMH's VT? This has actually been documented in another dynamic (MRI) video published by the Freiburg Institute of Musician's Medicine, *Naturtonreihe in Zungentechnik* (see youtu.be/-jKl61Xxkh0). It was taken when AMH performed overtone singing, enhancing, one by one each overtone of a drone with a fundamental frequency of 270 Hz (pitch about C4), in a rising followed by a descending sequence. Henceforth, the frequencies of the enhanced overtones will be referred to as F_E. All overtones, from the 4th, F_E = 1,080 Hz, up to the 12th, $F_E \approx$ 3,200 Hz, were enhanced.

The MRI video shows her entire VT in a midsagittal lateral profile. **Figure 6** shows tracings of the VT for each of the enhanced overtones in the ascending and the descending series.

Voice Source

Is formant clustering an exhaustive explanation of overtone singing? Fortunately, the transfer function of the VT can be predicted given its formant frequencies. Thus, a vowel spectrum can be analyzed not only with respect to the formant frequencies, which appear as peaks in the spectrum envelope, but also with respect to the voice source. The trick is simple, inverse filtering!

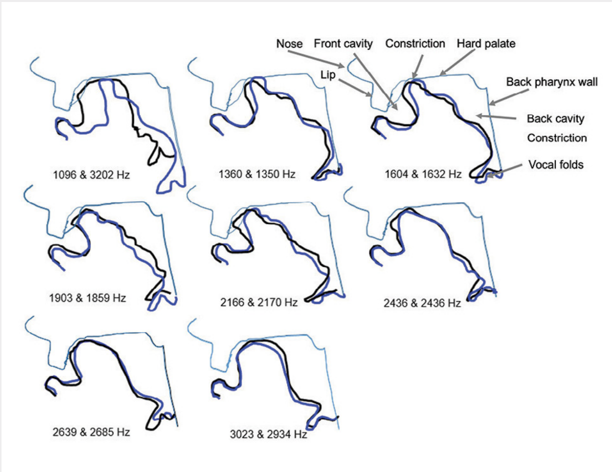


Figure 6. *Tracings of the lateral midsagittal articulatory profiles observed in the MR images of the subject while producing the indicated values of* F_E *in the ascending (**blue**) and descending (**black**) sequences.*

On its way from the glottis to the lip opening, the voice source has been filtered by the transfer function of the VT. Inverse filtering means that the radiated spectrum is filtered by the VT transfer function (**Figure 2C**) turned upside down.

The transfer function itself can be computed from the formant frequencies. This may sound a bit circular, but it is not. Glottal airflow must be zero when the glottis is closed. Hence the closed phase of the voice source waveform contains no ringing at a formant frequency if the inverse filter exactly equals the transfer function. Moreover, it is well-known that the spectrum envelope of the voice source has a smooth spectrum envelope, so peaks and valleys close to the formant frequencies are signs of inaccurate tuning of the inverse filters. Thus fine tuning of the inverse filters is a condition for reaching an accurate result. Errors reveal themselves in terms of ringing during the closed phase and/or a spectrum envelope peak and/or a trough near the formants.

The voice source can be varied along three dimensions. By stretching and tensing the vocal folds, the fundamental frequency increases, resulting in an increase in pitch. By increasing the overpressure of air in the respiratory system, the amplitude of the voice source increases, which causes vocal loudness to increase. By changing vocal fold adduction, which results in squeezing the glottis, the voice timbre varies along a dimension that ranges from breathy to pressed. Breathy phonation is what you typically use during a concert when you want to tell something to the person sitting next to you without disturbing the performance. Pressed phonation is typically used when you speak in excited anger or when you attempt to say something when carrying something very heavy.

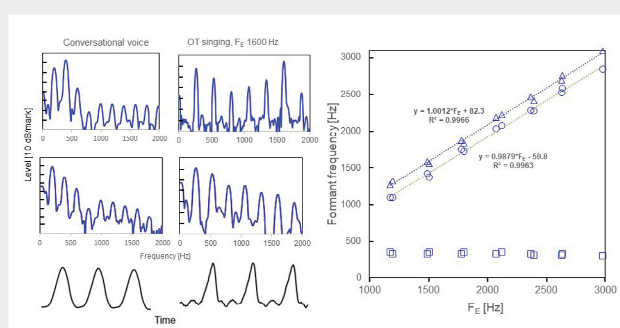
Figure 7 shows two examples of AMH's voice. Compared with her conversational speech, the source spectrum envelope slopes less steeply in overtone singing. Furthermore, the waveform has a longer closed phase and contains sharp knees, typical signs of an increase of vocal fold adduction. The ripple during the closed phase on overtone singing does not correspond to a formant frequency, but to a 900-Hz periodicity, an artifact frequently observed in glottal flow waveforms.

The frequencies of the three lowest formants used for the inverse filtering are plotted as function of F_E in **Figure 7**, *right*. The trend lines show that F2 and F3 have similar slopes and intercepts differing by about 185 Hz. Thus F2 and F3 are closely clustered around F_E, suggesting an affirmative answer to the question raised above, if formant clustering is the sole explanation of overtone singing. As formant frequencies are controlled by the shape of the VT, the next question then is how AMH shapes her VT to achieve this distribution of formant frequencies.

Estimating Vocal Tract Shapes

The resonances of the VT are determined by its shape, and we have excellent tools for varying this shape within

Figure 7. *Examples of AMH's voice in conversational speech (**left**) and during overtone (OT) singing (**center**). **Top:** radiated spectra. **Center:** voice source spectra; **Bottom:** glottal airflow waveforms. **Right:** the three lowest formant frequencies used for the inverse filtering of AMH's overtone singing as functions of* F_E*. **Lines** and **equations** represent trend lines.*



very wide limits. We can vary the shape of the tongue body, the position of the tongue tip, the jaw and lip openings, the larynx height, and the position and status of the gateway to the nose, the velum.

Let us now more closely examine the shape of AMH's VT as documented in the MRI video. It is evident from **Figure 6** that AMH produced overtone singing with a lifted tongue tip, so the tongue tip divided the VT into a front cavity and a back cavity. Our first target is the back cavity posterior to the raised tongue tip.

The formant frequencies associated with a given VT shape can be estimated from the VT contour. Several investigations have examined the relationship between the sagittal distance separating the VT contours and the associated cross-sectional area at the various positions along the VT length axis (see, e.g., Ericsdotter, 2005). Hence it was possible to describe the shape of the back cavity for each F_E in terms of an area function that lists the cross-sectional area as a function of the distance to the vocal folds.

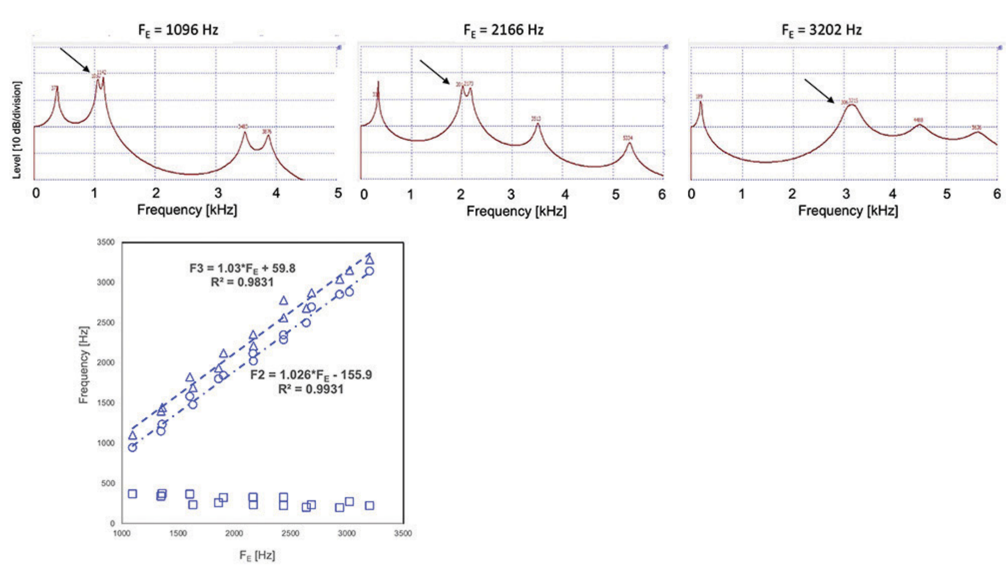
The next question concerns the front cavity, anterior to the raised tongue tip. The cavity between palatal constriction and the lip opening looks like, and can be regarded as, a Helmholtz resonator, a cavity in front of the tongue tip and a neck formed by the rather narrow lip opening (Sundberg and Lindblom, 1990; Granqvist et al., 2003).

The area of the lip opening was measured in a front video recorded when AMH produced the same overtone series as for the MRI video. The length of the lip opening was documented in the MRI video. These measures plus the frequency of the third formant used for the inverse filtering analysis allowed us to use the Helmholtz equation for calculating the front cavity volume. The validity of this approximation was corroborated in terms of a strong correlation between the measured length and the volume of the front cavity.

The formant frequencies of the entire VT could be calculated by a custom-made software, *Wormfrek* (Liljencrants and Fant, 1975). **Figure 8** shows the transfer functions with the formant frequencies for three F_E values: 1,096, 2,166, and 3,202 Hz. In **Figure 8**, the *arrows* highlight the close proximity of F2 and F3. In **Figure 8**, *bottom*, F1, F2, and F3 are plotted as a function of F_E. The trend lines show that F2 and F3 have similar slopes and intercepts differing by about 220 Hz.

We note that the F1, F2, and F3 predictions parallel the formant measurements made using inverse filtering (**Figure 7**). Here, a somewhat wider distance separates F2 from F3 than what was shown in **Figure 7**. The common

Figure 8. ***Top:*** *Wormfrek software displays of the transfer functions for the lowest, a middle, and the highest* F_E *(**left, center,** and **right,** respectively).* ***Bottom:*** *associated values of F1, F2, and F3 as a function of* F_E*.* ***Lines*** *and* ***equations*** *refer to trend line approximations.*



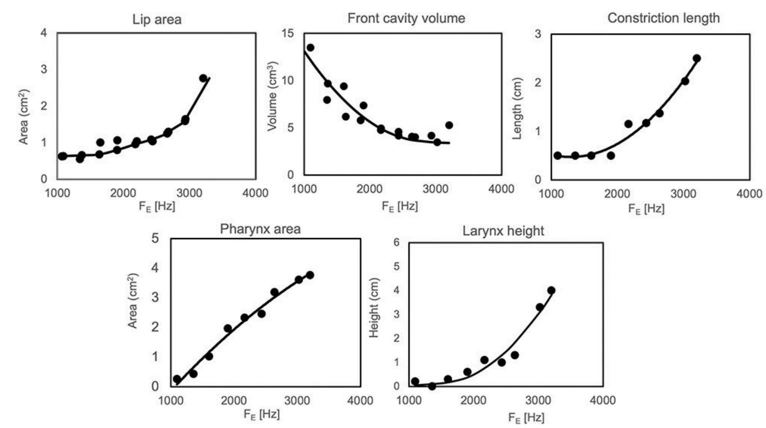


Figure 9. *Articulatory area function parameters plotted as functions of F_E. **Curves** show approximations derived from trend line equations.*

denominator is the consistent identification of the double formant. We feel justified in concluding that our results confirm the double formant phenomenon as a prerequisite for the overtone selection and enhancement in AMH's overtone singing technique.

Conclusions

Central to the present account is the "double formant" hypothesis, which attributes the phenomenon of overtone singing to VT filtering. However, the inverse filtering results also suggest that overtone singing involves a phonation type different from that in conversational voice, making the source spectrum slope less steep and thus boosting the amplitudes of the higher overtones. These findings replicate and extend previous investigations of overtone singing. Bloothooft et al. (1992) undertook an acoustic study of an experienced overtone singer and suggested formant clustering as an explanation and also noted an extended closed phase of the vocal fold vibrations. Using impedance measurements, Kob (2004) analyzed a form of overtone singing called sygyt and interpreted the overtone boosting as the result of formant clustering.

Parallel vibrations of the ventricular folds have been documented in throat singing (Lindestad et al., 2001). How about this possibility in AMH's overtone singing? Our inverse filtering data clearly rule out the existence of a laryngeal mechanism that selectively amplifies and enhances individual partials.

Overtone singing clearly requires an extremely high degree of articulatory precision; for each F_E, two cavities need to be shaped such that they produce resonance frequencies that match each other within a few tens of Hertz. How can the underlying motor control be organized? It is probably relevant that some of the articulatory configurations shown in **Figure 6** are used also in speech. The lateral profile for F_E = 1,096 Hz resembles the articulation of retroflex consonants (Dixit 1990; Krull and Lindblom, 1996). A narrow pharyngeal constriction is typical of [a]-like vowels and pharyngealized consonants (Ladefoged and Maddieson, 1996). The VT for F_E = 3,202 Hz has a "palatalized' tongue shape similar to that used for the vowel [i].

It would also be relevant that the articulatory parameters varied systematically with F_E. This is illustrated in **Figure 9**. It shows how AMH varied the lip opening area, length of palatal constriction, larynx height, front cavity volume, and pharynx area as a function of F_E. It is evident that the values of each individual articulatory dimension are aligned along smooth contours running between its values in F_E = 1,096 and 3,202 Hz. This lawful patterning suggests that it would be possible to derive VT shapes intermediate between those for F_E = 1,096 and 3,202 Hz by interpolation. A rough description would be to say that the VT shapes are located along a trajectory in the articulatory space that runs between a retroflex and pharyngealized [a] and an [i]-like, palatalized tongue profile.

How to Learn Overtone Singing

Producing overtones with your own voice is relatively easy. You practice singing very slow vowel transitions between the vowels /i/ and /u/ on a long-sustained drone that is kept at a constant pitch. Then, overtones start to appear quite clearly from your voice, although you might not able to hear them yet. To hear overtones in your own voice is the key to achieving deliberate control; learning to hear them is the first important part of your practice.

In this article, we have analyzed an advanced technique of overtone singing, double resonator articulation. The tongue tip is retracted and elevated in the mouth as for the American consonant /r/. This lowers the third formant and can bring it close to the second formant. As we have seen, this creates a double resonator and a double formant, which results in a strong, whistling-like overtone. To do this requires quite an accurate and simultaneous control over the front cavity for the third formant and the back cavity for the second formant. Generally, it takes quite some practice to learn this technique.

A simpler start into the fascinating world of overtone singing may be to learn to enhance overtones with vowels only, with an undivided VT cavity. Then, the VT works as a single resonator, and the second formant is solely responsible for overtone enhancement. Also, this technique can be learned by very slowly changing the articulation between /i/ and /u/, keeping a drone with constant pitch. When you manage to do this, you will discern single overtones; one by one, they first increase and then decrease in loudness as they approach the second formant, pass it, and then move away from it, and soon after, the next overtone will appear and do the same thing.

After you have learned the vowel-technique well, it is mostly both exciting and not too difficult to learn the double formant technique. Then, you may want to explore the pleasure of shifting the drone pitch and so extend the melodic possibilities of overtone singing even further. If you want to learn more, see Hefele (2020)!

Acknowledgments

Bernhard Richter, Matthias Echternach, Louisa Traser, and Michael Burdumy at the Freiburg Institute of Musician's Medicine are acknowledged for their work with the two dynamic magnetic resonance imaging videos.

References

Bloothooft, G., Bringmann, E., van Cappellen, M., van Luipen, J. B., and Thomassen, K. P. (1992). Acoustics and perception of overtone singing. *The Journal of the Acoustical Society of America* 92, 1827-1836.

Dixit, R. P. (1990). Lingotectal contact patterns in the dental and retroflex stops of Hindi, *Journal of Phonetics* 11, 291-302.

Ericsdotter, C. (2005). *Articulatory-Acoustic Relationships in Swedish Vowel Sounds*, Doctoral Dissertation, Stockholm University, Stockholm, Sweden.

Fant, G. (1960). *The Acoustic Theory of Speech Production*, Mouton, The Hague, The Netherlands.

Fant, G. (1972). Vocal tract wall effects, losses and resonance bandwidths. *Quarterly Progress and Status Report* 2(3), Department of Speech Hearing and Music (KTH), Stockholm, Sweden, pp. 28-52. Available at https://rb.gy/ugtss4.

Gill, K. Z., and Purves, D. (2009). A biological rationale for musical scales. *PLoS ONE* 4(12): e8144.

Granqvist, S., Sundberg, J., Cortes, E. E., Larsson, J., and Branderud, P. (2003). The front and sub-lingual cavities in coronal stops: An acoustic approach to volume estimation. *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona, Spain, August 3-9, 2003, pp. 941-944.

Hefele, A.-M. (2020). Anna-Maria Hefele's Website www.overtone.academy, Available at: https://rb.gy/yplwko. Accessed: October 12, 2020.

Kob, M. (2004). Analysis and modelling of overtone singing in the sygyt style. *Applied Acoustics* 65 1249-1259.

Krull, D., and Lindblom, B. (1996). Coarticulation in apical consonants: acoustic and articulatory analyses of Hindi, Swedish and Tamil, *Quarterly Progress and Status Report 2*, Department of Speech Hearing and Music (KTH), Stockholm, Sweden, pp. 73-76. Available at https://rb.gy/ugtss4.

Ladefoged, P., and Maddieson, I. (1986). *The Sounds of the World's Languages*, Blackwell, Oxford, UK.

Liljencrants, J., and Fant, G. (1975). Computer program for VT-resonance frequency calculations, *Quarterly Progress and Status Report 4*, Department of Speech Hearing and Music (KTH), Stockholm, Sweden, pp. 15-20. Available at https://rb.gy/ugtss4.

Lindestad, P.-Å., Södersten, M., Merker, B., and Granqvist, S. (2001). Voice source characteristics in Mongolian "throat singing" studied with high-speed imaging technique, acoustic spectra, and inverse filtering, *Journal of Voice* 15, 78-85.

Moore, B. C. J. (2012). *An Introduction to the Psychology of Hearing*, Emerald Group Publishing Ltd,. Bingley, UK

Smith, H., Stevens, K., and Tomlinson, R. (1967). On an unusual mode of singing by certain Tibetan lamas. *The Journal of the Acoustical Society of America* 41, 1262-1264.

Sundberg, J., and Lindblom, B. (1990). Acoustic estimation of the front cavity in apical stops *The Journal of the Acoustical Society of America* 88, 1313-1317.

Wendler J., Dejonckere, P. H., Smit, P. H., and Cebulla, M. (1991). Overtone singing in the view of Manuel Garcia. *Proceedings of Conference "New Ways of the Voice,"* Service ORL et Audiophonologie, Besancon, France, February 14-17, 1991, p. 46.

About the Authors

Johan Sundberg jsu@kth.se

Department of Speech Music and Hearing
Royal Institute of Technology (KTH)
SE-100 44 Stockholm, Sweden

Johan Sundberg studied musicology at Uppsala University (Uppsala, Sweden). He made his dissertation research on the acoustics of organ pipes as a guest researcher in Gunnar Fant's department at the Royal Institute of Technology (KTH; Stockholm, Sweden). This brought him into a productive contact with Björn Lindblom. After finishing his dissertation in 1966, he founded a research group in the area of music acoustics at the KTH and was awarded a personal chair in music acoustics there in 1979. Being an active singer, the voice as a musical instrument has been his main research theme along with the theory underlying music performance.

Author photo by Linnéa Heinerborg

Björn Lindblom
lindblom@ling.su.se

Department of Linguistics
Stockholm University
SE-106 91 Stockholm, Sweden

Björn Lindblom became an experimental phonetician in the early 1960s. His publications span a wide range of topics, including the development, production, and perception of speech. Academic experience: teaching and doing laboratory research at the Royal Institute of Technology (KTH; Stockholm, Sweden), Haskins Laboratories (New Haven, CT), MIT (Cambridge, MA), Stockholm University (SU), and the University of Texas at Austin (UT). He has held endowed chairs at SU and UT. He is a Fellow of the Acoustical Society of America and of the American Association for the Advancement of Science (AAAS) and is an Honorary Life Member of the Linguistic Society of America (LSA). His current project is a book: *Reinventing Spoken Language — The Biological Way*.

Anna-Maria Hefele
ama@anna-maria-hefele.com

Overtone Academy
Saulengrainerstrasse 1
DE-87742 Dirlewang, Germany

Anna-Maria Hefele has a Master of Arts from Mozarteum Salzburg (Austria) and is a multi-instrumentalist singer and overtone singer, performing worldwide as a soloist with different ensembles, choirs, and orchestras. She frequently performs in contemporary ballet, circus, and dance theater productions. Her YouTube video "Polyphonic Overtone Singing" went viral and has resulted in more than 17 million views so far, followed by regular appearances in various international television shows and radio broadcasts. Headlines like "A Voice as from Another World," "The Lady with the Two Voices," and "Polyphonic Vocalist Does the Impossible" have spread across the world.

Author photo by Thomas Radlwimmer

Global Positioning Systems: Over Land and Under Sea

Lora J. Van Uffelen

Introduction

Imagine with me a pre-COVID world. We are at an Acoustical Society of America (ASA) meeting in, say, Chicago, IL. We've just enjoyed a stimulating afternoon session, and our brains are fried. We need to find a coffee shop for a chat and some caffeine. What's the first thing we do? We quickly pull out our mobile phones and open a Yelp or Google Maps app to find a location within a five-minute walk of the conference venue with a four- or five-star review, and we are on our way, following turn-by-turn directions until we reach the destination. This mapping solution is delivered courtesy of a Global Navigation Satellite System (GNSS).

It is hard to imagine a world without GNSS. Even during quarantine, when we cannot be out having coffee with our colleagues in a new and exciting city, the motivated among us still use mapping applications to chart out our neighborhood walk or bike ride to see how far we have gone. We can "drop a pin" or share our location with the push of a button and find a friend in a parking lot or in the middle of the woods.

This positioning has become indispensable in the land, air, and space domains; however, as the electromagnetic signals sent by satellite systems do not transmit well in water, they are not available for undersea applications. Acoustic signals, however, propagate very well underwater and are commonly used for navigation of underwater vehicles as well as tracking marine mammals, fishes, turtles, and even lobsters. Typically, this tracking is done at short propagation ranges, but long-range signals can be used for positioning as well. Would it be possible to have a "Global Navigation Acoustic System" for the underwater domain that would be an analogue to the GNSS that we have become so reliant on? To answer this question, let's first familiarize ourselves with the GNSS.

Global Navigation Satellite Systems

Positioning from the GNSS is available all over the globe to provide localization, tracking, navigation, mapping, and timing, all of which are closely related but separate applications. Their availability and use have transformed the world in which we live. The most obvious relevancy of a GNSS is for the transportation industry. Mapping and route-planning applications include traffic avoidance features that have saved millions of dollars, reduced emissions, and limited time wasted in traffic. Aircraft pilots rely on the GNSS among other instruments when visual observations are not reliable.

Even the agriculture industry has been revolutionized by the GNSS. In the current age of precision agriculture, positioning systems in tractors and farm equipment can have centimeter accuracy to ensure that all of the fields are covered without driving over the same area twice. Snowplow operators also use the GNSS to locate edges of roads covered in snow.

Scientists who do field work, whether on land or at sea, would be lost without the GNSS. We rely on these satellite systems to locate our sensors and associate data with a position on the earth. Metadata for any type of dataset, acoustic or otherwise, typically contains time and location data provided by the GNSS.

The modern military uses the GNSS for guided missiles and drones to minimize collateral damage. In fact, the Global Positioning System (GPS), the US-based GNSS that we used to find coffee on our hypothetical trip to Chicago, was originally developed for military purposes. Before the development of the GPS, it was the task of several soldiers, sailors, or pilots to navigate the troops, ships, or planes. Tasking this to a remote and automated system minimizes the number of people involved in the

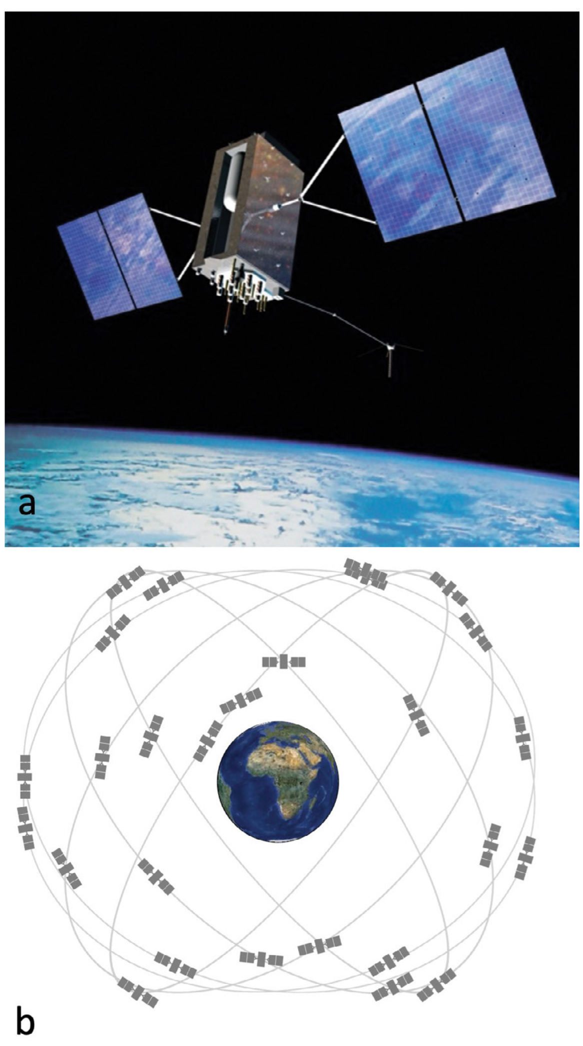
https://doi.org/10.1121/AT.2021.17.1.52

operations, freeing them up for other tasks as well as reducing human errors.

Overview and History of the Global Positioning System

The GPS is owned by the US Government and is operated and maintained by the US Air Force out of Shriever Air Force Base in Colorado. The system is young, relatively speaking. The GPS project was started in 1973, and the first NAVigation System with Timing and Ranging (NAVSTAR) satellite was launched in 1978. The 24-satellite system became fully operational in 1995. A photograph of a modern GPS-III satellite and a depiction of the satellite constellation are shown in **Figure 1**.

Figure 1. ***a:*** *Image of Global Positioning System (GPS) III satellite. A GPS III satellite is roughly the size of a small car and orbits approximately 20,200 km above the earth.* ***b:*** *Configuration of satellite constellation. The original constellation contained 6 orbitals with slots for 4 satellites each, and only 24 satellites are required to operate at any given time. But, in 2011, this was expanded to accommodate additional satellites to improve coverage. Source: United States Government (available at gps.gov/multimedia/images).*



A precursor to the GPS, the Transit System, was designed specifically to provide accurate location information to US Navy Polaris nuclear submarines and became the first operational satellite navigation system in 1964 (Guier and Weiffenbach, 1998). There is not enough space here to go into all of the science and technology advances that paved the way for the modern GPS, including the satellite geodesy work of Gladys Mae West (featured in Shetterly, 2016, and the movie *Hidden Figures*), but the gps.gov website, maintained by the national coordination office for space-based positioning, navigation, and timing and hosted by the National Oceanic and Atmospheric Administration (NOAA) has a wealth of useful information and links.

The US-based GPS satellites are not the only navigational satellites orbiting the earth. Indeed, a Russian-based GLObal NAVigation Satellite System (GLONASS) became operational around the same time as the GPS. The United States and Russia both started construction of their own GNSS constellations at the height of the Cold War. More recently, in June 2020, China launched the final satellite in the third generation of the BeiDou Navigation Satellite System, which now provides worldwide coverage. Europe has launched Galileo, which has been operational since 2019, and the complete 30-satellite system is expected by the time you read this article. Galileo is the only purely civilian system; the systems launched by the United States, China, and Russia are all at least partially owned or operated by the military. Most modern smartphones have the capability to receive signals from multiple constellations.

How Does the Global Positioning System Work?

GPS satellites continuously broadcast electromagnetic signals that travel through the atmosphere, providing their location and precise timing information. (I refer to the GPS, but this can be applied more globally to other GNSS constellations as they operate using the same principles.) A GPS satellite transmits multiple signals,

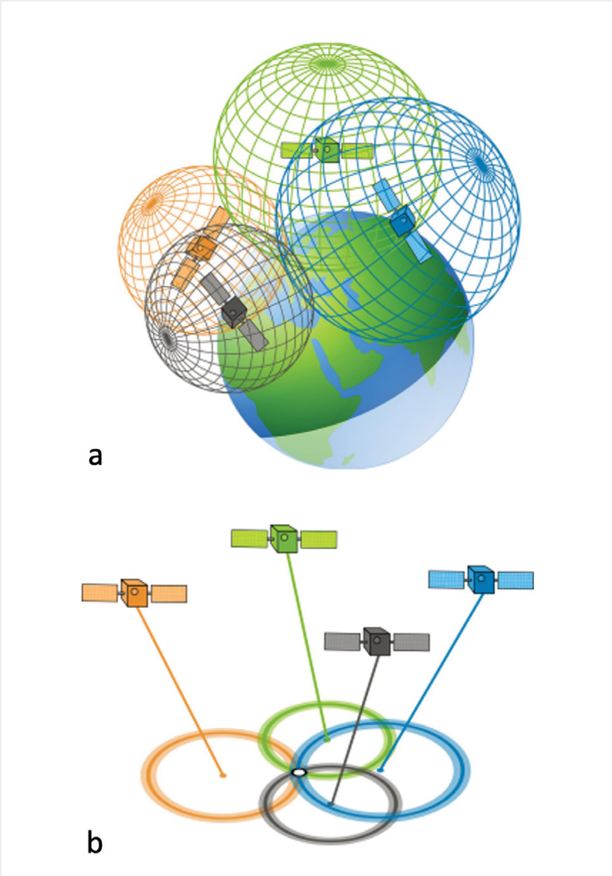


Figure 2. ***a:*** *Depiction of positioning of a GPS receiver using trilateration in three dimensions. Each sphere with a satellite at its center represents the distance calculated from Eq. 1. The four spheres intersect at the location of the GPS receiver.* ***b:*** *Trilateration with four satellites projected onto two dimensions with the single point of intersection determining the location of the GPS receiver. Presented with permission from gisgeography.com.*

including ranging signals and navigation messages. The original GPS design contained two ranging signals: a coarse/acquisition (C/A) code and a restricted precision (P) code reserved for military and government applications. Each satellite transmits a C/A code with a carrier frequency of 1,575.42 MHz. Galileo and BeiDou also transmit signals at this carrier frequency, which is in the microwave band, outside of the visible spectrum.

The time that it takes the signal to reach the receiver is used to calculate a range or distance (d) from the satellite with the following simple relationship

$$d = c \times t \tag{1}$$

where c is the speed of light (299,792,458 m/s) and t is the time of flight for the signal traveling through space. This time of flight is the difference between the time the signal is broadcast by the satellite and the time the signal is received. Once a GPS receiver obtains the distance between itself and at least four satellites, it can use geometry to determine its location and simultaneously correct its time.

The concept is relatively simple and is demonstrated in **Figure 2**. If we know just the range from a single satellite, the location of the receiver could be anywhere on an imaginary sphere with the satellite located at its center. Combining the ranges received from 4 satellites, along with precise timing information, provides a single intersection point in 3-dimensional space that corresponds to the position of the GPS receiver. This is referred to as trilateration (often confused with triangulation, which involves measuring angles rather than distances).

Apparent from **Eq. 1**, an inaccurate estimate of the signal travel time will give an incorrect distance from receiver to satellite and therefore an inaccurate position. Precise timing is therefore vital to GPS operation. Nanoseconds in timing error on the satellites lead to meters of positioning error on the ground. Each satellite has an atomic clock onboard, which provides precise timing information. These precise clocks are updated twice a day to correct the clock's natural drift using an even higher precision atomic clock based on land.

Underwater Positioning and Navigation Using Acoustics

It is interesting to note that satellite navigation systems were first designed with submarines in mind even though the GPS is not useful beneath the sea surface. Electromagnetic waves from the satellites travel very efficiently through the atmosphere but are quickly attenuated underwater. Underwater vehicles and underwater instrumentation are therefore unable to take full advantage of the GPS infrastructure.

Submarines do take advantage of underwater acoustic signals, and the field of underwater acoustics has largely been driven by military applications (Muir and Bradley, 2016). Acoustic waves are mechanical pressure waves and therefore do not propagate well in the (near) vacuum of space but travel more efficiently and more quickly in denser media. Because of this, sound travels faster in seawater than in air and it is less quickly attenuated.

The same basic relationship from **Eq. 1** that is used to calculate the distance from satellites can be applied to acoustic signals as well. Here, rather than multiplying the time that the GPS signal has traveled by the speed of light, the travel time of the signal is multiplied by the speed of sound in the medium through which it is traveling. The speed of sound in the ocean is roughly 1,500 m/s. This is much slower than the speed of light, and it is also quite variable because the speed of sound in seawater depends on the seawater temperature, salinity, and depth.

Traditional Underwater Positioning and Local Vehicle Navigation Systems

Underwater vehicles routinely get position and timing from a GPS receiver when they are at the surface, but once they start to descend, this is no longer available. Vehicles navigate underwater using some combination of dead reckoning, vehicle hydrodynamic models, inertial navigation systems (INSs), and local navigation networks (Paull et al., 2014). Positioning in the z direction, the depth in the ocean, is straightforward with a pressure sensor, which can reduce the dimensionality of the problem to horizontal positioning in x and y, or longitude and latitude, respectively.

Dead reckoning estimates the position using a known starting point that is updated with measurements of vehicle speed and heading as time progresses. Larger vehicles, such as submarines, may have an onboard INS that integrates measurements of acceleration to estimate velocity and thereby position. These measurements are, however, subject to large integration drift errors.

Because of the need for more position accuracy than afforded by the submarine systems discussed above, it comes as no surprise that underwater vehicles also use acoustics for localization. A long-baseline (LBL) acoustic-positioning system is composed of a network of acoustic transponders, often fixed on the seafloor with their positions accurately surveyed. The range measurements from multiple transponders are used to determine position. LBL systems typically operate on scales of 100 meters to several kilometers and have accuracies on the order of a meter. Transponder buoys at the surface can also provide positioning accuracy similar to a seafloor LBL network. These buoys have constant access to GPS positioning so they do not require a survey.

Short-baseline (SBL) systems operate on a smaller scale, and the SBL transducers are typically fixed to a surface vessel. Ultrashort-baseline (USBL) systems are typically a small transducer array, also often fixed to a surface vehicle, which use phase (arrival angle) information of the acoustic signals to determine the vehicle position.

These types of acoustic localization work in a similar way to GPS localization, with electromagnetic waves; however, they all operate in relatively small regions. Note that these acoustic-positioning methods have been described in the context of underwater vehicles, but they can be used for other purposes as well, including tracking drifting instrumentation or even animals underwater.

Long-Range Underwater Acoustics

Propagation in the SOFAR Channel

Attenuation of acoustic signals in the ocean is highly dependent on frequency. The signals commonly used for LBL, SBL, and USBL localization networks typically have frequencies of tens of kilohertz and upward. These signals may travel for a few kilometers, but lower frequency signals on the order of hundreds of hertz or lower are capable of traveling across entire ocean basins underwater. This was demonstrated in 1991 by the Heard Island Feasibility Test, where a signal was transmitted from Heard Island in the Southern Indian Ocean and received at listening stations across the globe, from Bermuda in the Atlantic Ocean to Monterey, CA, in the Eastern Pacific Ocean (Munk et al., 1994).

Refractive effects of the ocean waveguide are usually taken into account when using the acoustic-positioning methods described above because an acoustic arrival often does not take a direct path from the source to the receiver, and often a number of arrivals resulting from multiple propagation paths are received. The refractive effects of the ocean waveguide become even more important as ranges increase. Acoustic arrivals can be spread out over several seconds; however, the time arrival structure can be predicted based on the sound speed profile.

The speed of sound in the ocean increases with increasing hydrostatic pressure (depth in the ocean) and with higher temperatures that occur near the surface. This leads to a sound speed minimum referred to as the sound channel axis, which exists at approximately 1,000 m depth, although the depth can vary depending on where you are on the globe (**Figure 3**).

The SOFAR channel, short for SOund Fixing And Ranging, refers to a sound propagation channel (Worzel et al., 1948) that is centered around the sound channel axis. Sound from an acoustic source placed at the sound speed minimum will be refracted by the sound speed profile, preventing low-angle energy from interacting with the lossy seafloor and enabling the sound rays to travel for very long distances, up to thousands of kilometers.

The rays take different paths when traveling over these long ranges, as seen in **Figure 3**. The arrival time at a receiver is an integrated measurement of travel time along the path of the ray. Rays that are launched at angles near the horizontal stay very close to the sound speed minimum. Rays that are launched at higher angles travel through the upper ocean and deep ocean, and although they take a longer route than the lower angle rays, they travel through regions of the ocean that have a faster sound speed and therefore arrive at a receiver before their counterparts that took the shorter, slower road.

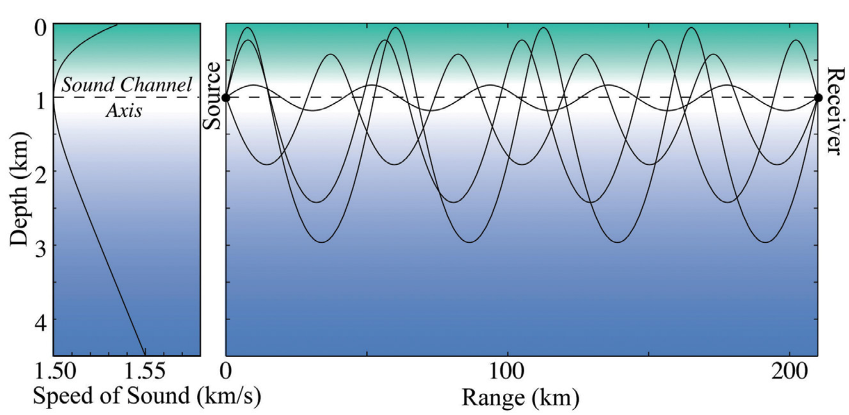
Ocean Acoustic Tomography Measurements

Ocean acoustic tomography takes advantage of the variability in measured travel times for specific rays to invert for ocean temperature. Each ray has traveled a unique path through the ocean and therefore carries with it information on the sound speed along the particular path that it has traveled. On a very basic level, we are looking again at the relationship from **Eq. 1**, but here distance and travel time are known, and we are inverting for sound speed, which is a proxy for temperature. In ocean acoustic tomography, the variability in these acoustic travel times is measured regularly over a long period of time (acoustic sources and receivers often remain deployed in the ocean for a year at a time) to track how the ocean temperature is changing. This method was described by Worcester et al. (2005) in the very first issue of *Acoustics Today* and more thoroughly in the book, *Ocean Acoustic Tomography*, by Munk et al. (1995).

The variability in these travel times is measured in milliseconds; therefore, as with a GNSS, the acoustic travel time measurements must be extremely precise. Great care is taken to use clocks with low drift rates and to correct for any measured clock drift at the end of an experiment.

The locations of the acoustic sources and receivers also must be accurate because inaccuracies in either position would lead to an inaccurate calculation of distance, which would impact the inversion for sound speed based on the simple relationship of **Eq. 1**. The sources and receivers used in typical ocean acoustic tomography applications are on subsurface ocean moorings, meaning that there

Figure 3. ***Left:*** *canonical profile of sound speed as a function of depth in the ocean (**solid line**). **Right:** refracted acoustic ray paths from a source at 1,000 m depth to a receiver at 1,000 m depth and at a range of 210 km. The Sound Channel Axis (**dashed line**) is located at the sound speed minimum at a depth of 1 km. Adapted by Discovery of Sound in the Sea (see dosits.org) from Munk et al., 1995, Figure 1.1, reproduced with permission.*



is an anchor on the seafloor with a wire stretched up to a buoy that sits below the surface to hold the line taut. The sources and hydrophone receivers are mounted on this line. Additional floatation is also mounted on the line to keep the mooring standing upright, but it is subject to ocean currents, so it moves around in a watch circle about the anchor position. An instrument at the top of a 5,000-m mooring could be swept several hundred meters from the latitude and longitude position of the anchor by ocean currents. A LBL array of acoustic transponders, as described in ***Traditional Underwater Positioning and Local Vehicle Navigation Systems***, is typically deployed around each mooring position to track the motion of the sources and receivers throughout the experiment to correct for the changes in distance between the sources and receivers.

Positioning with Long-Range Underwater Acoustic Measurements

The same core concepts of inferring distance from measurements of signal travel time that we see in GNSS and local underwater acoustic networks can also apply at long ranges. Neutrally buoyant oceanographic floats called swallow floats were equipped with acoustic pingers to be tracked by a nearby ship; these were adapted to take advantage of the deep sound channel and were subsequently known as SOFAR floats. The first SOFAR float was deployed in 1968 and was detected 846 km away (Rossby and Webb, 1970).

The SOFAR float signals were originally received by the SOund SUrveillance System (SOSUS) of listening stations operated by the US military. This system tracked more than just floats and enemy submarines. It also received acoustic signals from earthquakes, and there is a wonderful 43-day record of passively tracking of an individual blue whale, nicknamed Ol' Blue, as it took 3,200-km tour of the North Atlantic Ocean (Nishimura, 1994).

The existing listening system was convenient, but equipping each float with an acoustic source was technologically challenging and expensive. In the 1980s, the concept was flipped so that the float had the hydrophone receiver, and acoustic sources transmitted to the floats from known locations to estimate range to the float. The name was also flipped, and the floats are known as RAFOS, an anadrome for SOFAR (Rossby et al., 1986).

RAFOS sources have been useful to track floats in open water, but when there is sea ice present and the float is unable to get to the surface for a GPS position, underwater positioning becomes even more important. A recent study in the Weddell Gyre near Antarctica tracked 22 floats under ice that were unable to surface to obtain position from the GPS for eight months (Chamberlain et al., 2018).

Similar to RAFOS, a separate long-range navigation system in the Arctic used surface buoys to transmit GPS positions to floats and vehicles for under-ice ranging, with an accuracy of 40 m over 400-km ranges. This system operated at 900 Hz, with a programmable bandwidth from 25 to 100 Hz (Freitag et al., 2015).

RAFOS signals have a bandwidth of 1.6 Hz and therefore less time resolution than a more broadband source. **Figure 4**, **a** and **b**, contrast predictions of the arrival structure at a 1,145-km range for a RAFOS source with a broadband source having a bandwidth of 50 Hz. In both cases, the latest arriving energy is concentrated near the depth of the sound channel axis, corresponding to rays that stayed at depths with low sound speeds. The early arrivals are from rays that ventured into the higher speed regions of the sound speed profile (in **Figure 3**, ***dark blue*** and ***green***) and therefore also span more of the ocean depth. In both cases, we can see that the energy is spread over about 4 s, but the broadband source provides better resolution.

Figure 4, **c** and **d**, shows slices of these acoustic predictions at a 2,000 m depth. The broadband signal shown in **Figure 4d** exhibits sharp peaks in the arrival that can be identified with individual ray paths.

The increased bandwidth is one of the design suggestions for a potential joint navigation/thermometry system addressed in Duda et al. (2006). A system of sources is suggested with center frequencies on the order of 100-200 Hz and a 50-Hz bandwidth.

The acoustic sources used for ocean acoustic tomography applications are broadband sources designed to transmit over ocean basin scales. A 2010-2011 ocean acoustic tomography experiment performed in the Philippine Sea featured six acoustic sources in a pentagon arrangement and provided a rich dataset for evaluating long-range positioning algorithms. The sources used in this particular experiment had a center frequency of about 250 Hz and a bandwidth of 100 Hz.

The sources were used to localize autonomous underwater vehicles that had access to a GPS at the sea surface but only surfaced a few times a day. Hydrophones on the vehicles received acoustic transmissions from the moored sources at ranges up to 700 km, and these signals were used to estimate the position of the vehicle when it was underwater (Van Uffelen et al., 2013). The measured acoustic arrivals were similar to the modeled arrival shown in **Figure 4d**. The measurements of these peaks collected on the vehicle were matched to predicted ray arrivals to determine range. This method takes advantage of the multipath arrivals in addition to signal travel time. As with other acoustic methods and with the GPS, ranges from multiple sources were combined to obtain estimates of vehicle position. The resulting positions had estimated uncertainties less than 100 m root mean square (Van Uffelen et al., 2015).

Other long-range acoustic-ranging methods incorporate predictions of acoustic arrivals based on ocean state estimates (Wu et al., 2019). An algorithm introduced by Mikhalevsky et al. (2020) provides a "cold start' capability that does not require an initial estimate of the acoustic arrival and has positioning orders on the order of 60 m. These results were validated using hydrophone data with known positions that received the Philippine Sea source signals. As with the aforementioned method, this algorithm relies on the travel-time resolution afforded by the broadband source signals.

How Feasible Is a Global Navigation Acoustic System?

Because acoustic signals are able to propagate over extremely long ranges underwater, acoustics could provide an underwater analogue to the electromagnetic GNSS signals that are used for positioning in the land, air, and space domains. There are definite differences between using an underwater acoustic positioning system and a GNSS, however. GNSS satellites orbit the earth twice a day and transmit continuously. Acoustic sources do not need to be in orbit, but proper placement of the sources would enable propagation to most regions in the oceans of the world.

The far reach of underwater acoustic propagation is demonstrated by the International Monitoring System (IMS) operated by the Comprehensive Nuclear Test Ban Treaty Organization (CTBTO). The IMS monitors the globe for acoustic signatures of nuclear tests with only six underwater passive acoustic hydrophone monitoring stations worldwide. **Figure 5** shows the coverage of these few stations. Signals received on these hydroacoustic stations were used to localize an Argentinian submarine that was lost in 2017 using acoustic recordings of the explosion on IMS listening stations at ranges of 6,000 and 8,000 km from the site (Dall'Osto, 2019).

You may note that **Figure 5** does not show much coverage in the Arctic Ocean and that the sound speed structure is quite different at high latitudes because it does not have the warm surface that we see in **Figure 3**; however, long-range propagation has been demonstrated in the Arctic

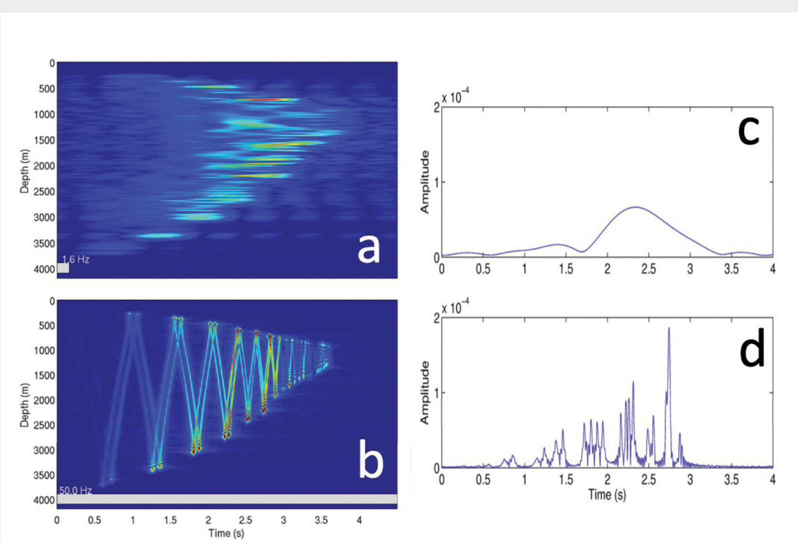


Figure 4. *Predictions of the acoustic arrival for a 260-Hz source at a range of 1,145 km. for a RAFOS source with a bandwidth of 1.6 Hz (**a**) and for a source with a bandwidth of 50 Hz (**b**). The arrivals in both cases are spread over about 4 s, with early arriving energy from higher angle rays and later arriving energy from rays launched at low angles that stayed near the depth of the sound channel axis. Slices of the plots shown in **a** and **b** were taken at a depth of 2,000 m for the RAFOS source (**c**) and broadband source (**d**) to contrast the travel time resolution. Adapted from Duda et al., 2006, with permission.*

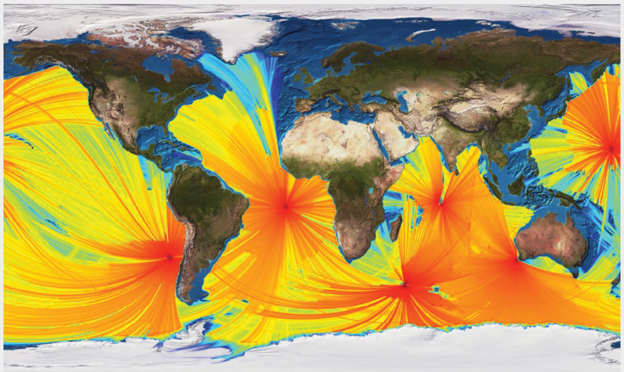


Figure 5. *Global coverage of the Comprehensive Nuclear Test Ban Treaty Organization (CTBTO) International Monitoring System (IMS), shown by a 3-dimensional model of low-frequency (<50-Hz) propagation. The property of reciprocity is invoked by placing sources at the locations of the six IMS hydrophone listening stations (**red areas**) from where the sound radiates. Colors represent transmission loss with a range of 70 dB. Figure created by Kevin Heaney and reproduced from Heaney and Eller, 2019, with permission.*

Ocean as well. In a 2019–2020 experiment, 35-Hz signals were transmitted across the Arctic Ocean over the North Pole (Worcester et al., 2020).

The electromagnetic signals broadcast by GNSS satellites are outside the visible spectrum, so we do not notice the signals that are continuously emitted by the satellites. In addition to the engineering challenges that would face continuous acoustic transmission, the frequency band of long-range propagation is within the hearing range of many animals, and the impacts to the environment, including potentially masking marine mammal vocalizations, would need to be considered. Long-range acoustic transmissions for scientific purposes go through an intense permitting process that takes into account the environment and the impacts on marine animals in the environment.

Each GNSS satellite broadcasts navigation messages that includes the date and time as well as the status of the satellite. It broadcasts ephemeris data that provide its specific orbital information for more precise localization of the GPS receiver. Localization using dedicated networks of sources, such as the example in the Philippine Sea, which incorporates precise source position and timing as necessary for localization of an acoustic receiver as it is for GPS has been discussed. A vision for a multipurpose acoustic observing system (Howe et al., 2019), would transmit this information as well to enable mobile platform positioning and navigation. Such a system could also provide ocean acoustic tomography measurements and passive acoustic monitoring for biological, natural, and anthropogenic sources.

Final Thoughts

The GPS satellite constellation was originally designed to meet national defense, homeland security, civil, commercial, and scientific needs in the air, in the sea, and on land. The age of artificial intelligence and big data has made GPS data on land incredibly useful to all of us in our everyday life. Not only can we use information on our own location from our cell phone to find the nearest coffee shop, we can take advantage of the location information on many different devices to look at traffic patterns to gauge what is the best way to get to that coffee shop. It won't be too long until we will be riding in self-driving cars, automatically taking the best route and precisely positioned relative to each other. All of this happened in just the last few decades because it has been only 25 years since GPS became fully operational.

An underwater analogue to a global navigation satellite system would revolutionize any operations in the underwater domain including oceanographic science, naval military applications, underwater vehicles, and even scuba diving. Acoustics is the most promising way to approach this on a large scale.

Acknowledgments

I extend my gratitude to Arthur Popper, Kathleen Wage, and Peter Worcester for their helpful suggestions and acknowledge the Office of Naval Research (ONR) for supporting my work related to underwater acoustic positioning.

References

Chamberlain, P. M., Talley, L. D., Mazloff, M. R., Riser, S. C., Speer, K., Gray, A. R., and Schwartzman, A. (2018). Observing the ice-covered Weddell Gyre with profiling floats: Position uncertainties and correlation statistics. *Journal of Geophysical Research Oceans* 123, 8383-8410. https://doi.org/10.1029/2017JC012990.

Dall'Osto, D. R. (2019). Taking the pulse of our ocean world. *Acoustics Today* 15(4), 20-28.

Duda, T., Morozov, A., Howe, B., Brown, M., Speer, K., Lazarevich, P., Worcester, P., and Cornuelle, B. (2006). Evaluation of a long-range joint acoustic navigation/thermometry system. *Proceedings of OCEANS 2006*, Boston, MA, September 18-21, 2006, pp. 1-6.

Freitag, L., Ball, K., Partan, J., Koski, P., and Singh, S. (2015). Long range acoustic communications and navigation in the Arctic. *Proceedings of OCEANS 2015-MTS/IEEE*, Washington, DC, October 19-22, 2015, pp. 1-5.

Guier, W. H., and Weiffenbach, G. C. (1998). Genesis of satellite navigation. *Johns Hopkins APL Technical Digest* 19(1), 14-17.

Heaney, K. D., and Eller, A. I. (2019). Global soundscapes: Parabolic equation modeling and the CTBTO observing system. *The Journal of the Acoustical Society of America* 146(4) 2848. https://doi.org/10.1121/1.5136881.

Howe, B. M., Miksis-Olds, J., Rehm, E., Sagen, H., Worcester, P. F., and Haralabus, G. (2019). Observing the oceans acoustically. *Frontiers in Marine Science* 6, 426. https://doi.org/10.3389/fmars.2019.00426.

Mikhalevsky, P. N., Sperry, B. J., Woolfe, K. F., Dzieciuch, M. A., and Worcester, P. F. (2020). Deep ocean long range underwater navigation, *The Journal of the Acoustical Society of America* 147(4), 2365-2382. https://doi.org/10.1121/10.0001081.

Muir, T. G., and Bradley, D. L. (2016). Underwater acoustics: A brief historical overview through World War II. *Acoustics Today* 12(2), 40-48.

Munk, W., Worcester, P., and Wunsch, C. (1995). *Ocean Acoustic Tomography*. Cambridge University Press, New York, NY.

Munk, W. H., Spindel, R. C., Baggeroer, A., and Birdsall, T. G. (1994). The Heard Island Feasibility Test. *The Journal of the Acoustical Society of America* 96, 2330-2342.

Nishimura, C. E. (1994). Monitoring whales and earthquakes by using SOSUS. *1994 Naval Research Laboratory Review*, pp. 91-101.

Paull, L., Saeedi, S., Seto, M., and Li, H. (2014). AUV navigation and localization: A review. *IEEE Journal of Oceanic Engineering* 39, 131-149.

Rossby, T., and Webb, D. (1970). Observing abyssal motions by tracking Swallow floats in the SOFAR channel. *Deep Sea Research and Oceanographic Abstracts* 17(2), 359-365. https://doi.org/10.1016/0011-7471(70)90027-6.

Rossby, T., Dorson, D., and Fontaine, J. (1986). The RAFOS system. *Journal of Atmospheric and Oceanic Technology* 3, 672-679.

Shetterly, M. L. (2016). *Hidden Figures: The American Dream and the Untold Story of the Black Women Mathematicians Who Helped Win the Space Race*. William Morrow, New York, NY.

Van Uffelen, L. J., Howe, B. M. Nosal, E. M. Carter, G. S., Worcester, P. F., and Dzieciuch, M. A. (2015). Localization and subsurface position error estimation of gliders using broadband acoustic signals at long range. *Journal of Oceanic Engineering* 41(3) 501-508. https://doi.org/10.1109/joe.2015.2479016.

Van Uffelen, L. J., Nosal, E. M., Howe, B. M., Carter, G. S., Worcester, P. F., Dzieciuch, M. A., Heaney, K. D., Campbell, R. L., and Cross, P. S.. (2013). Estimating uncertainty in subsurface glider position using transmissions from fixed acoustic tomography sources. *The Journal of the Acoustical Society of America* 134, 3260-3271. https://doi.org/10.1121/1.4818841.

Worcester, P. F., Dzieciuch M. A., and Sagen, H. (2020). Ocean acoustics in the rapidly changing Arctic. *Acoustics Today* 16(1), 55-64.

Worcester, P. F., Munk, W. H., and Spindel, R. C. (2005). Acoustic remote sensing of ocean gyres. *Acoustics Today* 1(1), 11-17.

Worzel, J. L., Ewing, M., and Pekeris, C. L. (1948). Propagation of Sound in the Ocean. *Geological Society of America Memoirs* 27, Geological Society of America, New York, NY. https://doi.org/10.1130/MEM27.

Wu, M., Barmin, M. P., Andrew, R. K., Weichman, P. B., White, A. W., Lavely, E. M., Dzieciuch, M. A., Mercer, J. A., Worcester, P. F., and Ritzwoller, M. H. (2019). Deep water acoustic range estimation based on an ocean general circulation model: Application to PhilSea10 data. *The Journal of the Acoustical Society of America* 146(6), 4754-4773. https://doi.org/10.1121/1.5138606.

About the Author

Lora J. Van Uffelen
loravu@uri.edu

Department of Ocean Engineering
University of Rhode Island
Narragansett, Rhode Island 02882, USA

Lora J. Van Uffelen is an assistant professor in the Department of Ocean Engineering, University of Rhode Island (Narragansett), where she teaches undergraduate and graduate courses in underwater acoustics and leads the Ocean Platforms, Experiments, and Research in Acoustics (OPERA) Lab. She earned her PhD in oceanography from the Scripps Institution of Oceanography, University of California, San Diego (La Jolla). Her current research projects focus on long-range underwater acoustic propagation, Arctic acoustics, vehicle and marine mammal localization, and acoustic sensing on underwater vehicles. She has participated in more than 20 research cruises, with over 400 days at sea.

Recent Acoustical Society of America Awards and Prizes

Acoustics Today is pleased to present the names of the recipients of the various awards and prizes given out by the Acoustical Society of America. After the recipients are approved by the Executive Council of the Society at each semiannual meeting, their names are published in the next issue of *Acoustics Today*.

Congratulations to the following recipients of Acoustical Society of America medals, awards, prizes, and fellowships, who will be formally be recognized at the Spring 2021 Plenary Session. For more information on the accolades, please see acousticstoday.org/asa-awards, acousticalsociety.org/prizes, and acousticstoday.org/fellowships.

Gold Medal

James F. Lynch
(Woods Hole Oceanographic Institution, Woods Hole, MA)

Helmholtz-Rayleigh Interdisciplinary Silver Medal in Physical Acoustics and Engineering Acoustics

Michael R. Moldover
(National Institute of Standards and Technology [NIST], Gaithersburg, MD; retired 3/19)

R. Bruce Lindsay Award

Likun Zhang
(University of Mississippi, University)

Medwin Prize in Acoustics Education

Ana Širović
(Texas A&M University, College Station)

William and Christine Hartmann Prize in Auditory Neuroscience

Philip Joris
(Katholieke Universiteit Leuven [Catholic University of Leuven; KU Leuven], Leuven, Belgium)

Congratulations also to the following members who were elected Fellows in the Acoustical Society of America in the spring 2021.

- **Kathryn H. Arehart**
(University of Colorado at Boulder)
for contributions to the understanding of auditory perception, hearing loss, and hearing aids
- **Gregory Clement**
(US Food and Drug Administration, Silver Spring, MD)
for contributions to transcranial therapeutic ultrasound
- **Ewa Jacewicz**
(Ohio State University, Columbus)
for contributions to the understanding of spectral and temporal dynamics in speech acoustics and perception
- **Joan A. Sereno**
(University of Kansas, Lawrence)
for contributions to speech learning, perception, and production across individuals and languages
- **Brian D. Simpson**
(Air Force Research Laboratory, Dayton, OH)
for contributions to speech perception, spatial hearing, and the development of auditory displays
- **Pamela E. Souza**
(Northwestern University, Evanston, IL)
for advancing understanding of the factors that affect an individual's response to hearing aid signal processing
- **Daniel J. Tollin**
(University of Colorado School of Medicine, Aurora)
for multidisciplinary contributions linking acoustics, physiology, and behavior to the understanding of binaural hearing
- **Matthew W. Urban**
(Mayo Clinic, Rochester, MN)
for outstanding contributions to the field of ultrasonic assessment of biologic tissue properties

https://doi.org/10.1121/AT.2021.17.1.61

Ask an Acoustician: Zoi-Heleni Michalopoulou

Zoi-Heleni Michalopoulou and Micheal L. Dent

Meet Zoi-Heleni Michalopoulou

This "Ask an Acoustician" essay features Zoi-Heleni (Eliza) Michalopoulou from the Department of Mathematical Sciences, New Jersey Institute of Technology (NJIT; Newark). Eliza is a member of the Acoustical Oceanography, Acoustical Signal Processing, and Underwater Acoustics Technical Committees of the Acoustical Society of America (ASA). She is a Fellow of the ASA and has been a member of the College of Fellows, cochair of the Women in Acoustics Committee, and an associate editor for *The Journal of the Acoustical Society of America*. I will let Eliza tell you the rest in her own words.

A Conversation with Zoi-Heleni Michalopoulou, in Her Own Words

Tell us about your work.

I listen to the ocean! Employing both acoustic models and statistical signal processing, I conduct research in geoacoustic inversion. That is, I solve the inverse problem, estimating properties of the seabed with which the sound has interacted. The forward problem in ocean acoustics, in simple terms, entails the modeling of the sound that travels in the ocean using mathematical relationships and relying on physics; the models connect sound propagation to the location of the source that transmits the sound, the placement of the receivers where the sound is sensed, and ocean environment parameters such as properties of the sediments. The inverse problem, on the other hand, uses these forward models and recorded data to move backward, that is, to identify the properties that generated the measured sound.

Geoacoustic inversion is one aspect of the inverse problem. My interests extend to inversion for source detection and location as well, both of which are inherently tied with geoacoustic inversion. Knowing the propagation environment, determined to a large degree by the seabed structure obtained via geoacoustic inversion, allows us to better detect, identify, and localize sources of interest in the ocean, whether these are submarines or sound-producing marine life. These tasks are of paramount importance in antisubmarine warfare and the study of marine life.

I see my work as a combination of underwater acoustics, acoustic signal processing, and acoustical oceanography, reflected in the three ASA technical committees of which I am a member. Recently, I have delved into machine-learning methods, both for sediment characterization and source localization. I am fortunate to have colleagues who share experimental data with me, which facilitates the validation of my methods in real environments.

Describe your career path.

I was born in Athens, Greece. Often, when I was in elementary school, I would get together with my best friend and we would put together electrical circuits from a set that had been given to my brother as a present. I then attended Pierce College, the high school of the American College of Greece. I had the opportunity there to be exposed to a rich curriculum in liberal arts, science and mathematics, languages, and art. Math attracted me the most, and I decided early on that I wanted to do something with numbers.

As is often the case, I was told that girls are not made for math and that only made me more determined to pursue

https://doi.org/10.1121/AT.2021.17.1.62

a STEM career. I decided to study electrical engineering at the National Technical University of Athens because it was the most prestigious STEM program in my hometown; the electrical circuits of my childhood may have played a role in my decision! I enjoyed my studies and decided to continue for an MS in electrical engineering in the United States. My high-school years at the American College of Greece had prepared me for this wonderful adventure. I was fortunate to attend the MS program at Duke University (Durham, NC) where I met my advisor, Dimitri Alexandrou, who had a passion for anything that had to do with sound and the ocean. He inspired me, and not only did I complete my MS thesis in ocean acoustic signal processing, but I decided to move forward for a doctorate in the same area.

Research became a passion, which led to an academic career. I have been at NJIT ever since I graduated from Duke, starting as a research assistant professor/postdoctoral fellow in the Department of Mathematical Sciences. The environment was (and still is) full of energy and a great fit. Soon after I joined, an opening for an assistant professor position came up. It was an easy decision for me to apply and accept the offer that followed. I have been enjoying a fruitful career there ever since. I have had the fortune to meet at Acoustical Society conferences colleagues such as Ross Chapman, Alex Tolstoy, Jim Candy, Ed Sullivan, Ellen Livingston, Leon Sibul, Leon Cohen, and many others, who mentored me in my early years and to them I owe much of the satisfaction I have been drawing from my career.

What is a typical day for you?

I am a morning person, and my day starts very early; I am up at 5 a.m. with a cup of coffee, reading *The New York Times* on my computer. But, other than that, every day is different. Research, teaching, and administration all compete for time. I try to get a good few hours of uninterrupted research time before delving into class preparation, teaching, and administration. I have frequent meetings with my students that I look forward to because they often lead to fresh ideas and perspectives. I draw a firm line at around 6 p.m. Family and personal time start then unless deadlines are looming. Relaxed family dinners, classes at the Adult School of my town, and reading occupy my evenings.

How do you feel when experiments/projects do not work out the way you expected them to?

I sometimes get frustrated, but I try to take it as a learning experience. I look for the reason behind the failure of an idea. That usually leads to a new idea that is an alternative look at the problem I need to solve. And I try to remind myself that progress in research is not a linear process.

Do you feel like you have solved the work-life balance problem? Was it always this way?

Yes, as much as this is possible. I have a supportive family and a flexible working environment. Teaching courses at convenient times and having family members help with child care so that I could attend conferences helped me attain a satisfying combination of career and family life. Having a daughter who appreciated my work and enjoyed telling her friends about her mom searching for submarines was a bonus! The flexibility of my work allowed me to get involved in the community. I served as a volunteer for my daughter's Girl Scout troop and unit, and I also volunteered at her elementary school, mostly helping students with math and science. I managed never to miss my daughter's recitals, choir events, and soccer or volleyball games. I enjoyed travel and still do, attending many conferences, often with my husband and daughter, and I get to visit frequently my family in Greece, where I also enjoy collaborations with colleagues at the National Technical University of Athens.

What makes you a good acoustician?

I work in an applied mathematics and statistics department in a technological institute that enables me to have discussions and collaborations with researchers from multiple areas in the mathematical and physical sciences as well as engineering. I develop new ideas and a better understanding of acoustics problems after I become exposed to research advances in different disciplines. And I learn from my students.

How do you handle rejection?

I put aside negative reviews and revisit them a couple of weeks later. I carefully consider critique (sometimes I agree and sometimes not) and try to use it to develop new ideas or better arguments for my existing ones. I keep going.

What are you proudest of in your career?

It has been a privilege to have mentored numerous bright and talented young people, several of them from underrepresented groups in the sciences. I have had the pleasure of guiding several women in research projects, both during their graduate and undergraduate studies. I take great pride in their accomplishments during and after their time at NJIT. I follow their career paths and keep in touch; notes that they send me decorate my office. Similarly, I have found it rewarding to address middle- and high-school students and to inspire them (I hope!) about pursuing careers in STEM. On several occasions, students have approached me afterward, startled and excited about careers in math that they had never imagined.

And, of course, I am exceedingly proud of the bright 23-year-old woman that my husband and I have raised in parallel to our careers, who has often inspired me to work harder so that I could become a better role model for her and her peers.

What is the biggest mistake you've ever made?

Overthinking everything. Writing a paper or research proposal was sometimes a particularly lengthy endeavor. Should I include the last figure? How about adding one more reference? And how about this email I need to send? How will I convey my message? Once I realized it, I stopped it and became more efficient and effective.

What advice do you have for budding acousticians?

Enjoy the journey into a multifaceted field. Attend conferences and listen to talks from all areas of acoustics; seek collaborations and cross-fertilization. Take risks and explore new directions.

Have you ever experienced imposter syndrome? How did you deal with that if so?

Yes, I did, in the very beginning of my career as a faculty member. With the advice of a wonderful colleague and mentor, I realized that the first person I needed to persuade that I truly belonged in a challenging academic environment was myself. Everything followed smoothly after that.

What do you want to accomplish within the next 10 years or before retirement?

I plan to continue with all my activities: research, teaching, and administration. What I would particularly like to accomplish is the mentoring of more undergraduate students in research. There is a spark when undergraduates are exposed to research questions and asked to work alongside graduate students, postdocs, and faculty. Several are inspired to go on to graduate school and some continuing to work in acoustics. Others tell me that their research experience and participation in research teams in their undergraduate years enables them to work more effectively in groups in their jobs in industry. A worthwhile experience all around.

Bibliography

Frederick, C., Villar, S., and Michalopoulou, Z.-H. (2020). Seabed classification using physics-based modeling and machine learning *The Journal of the Acoustical Society of America* 148, 859-872.

Lin, T., and Michalopoulou, Z.-H. (2016). A direct method for the estimation of sediment sound speed with a horizontal array in shallow water. *IEEE Journal of Oceanic Engineering* 42, 208-218.

Michalopoulou, Z.-H., Pole, A., and Abdi, A. (2019). Bayesian coherent and incoherent matched-field localization and detection in the ocean. *The Journal of the Acoustical Society of America* 146, 4812-4820.

Piccolo, J., Haramuniz, G., and Michalopoulou, Z.-H. (2019). Geoacoustic inversion with generalized additive models. *The Journal of the Acoustical Society of America* 145, EL463-EL468.

Contact Information

Zoi-Heleni Michalopoulou michalop@njit.edu

Department of Mathematical Sciences
New Jersey Institute of Technology
618 Cullimore
Newark, New Jersey 07102, USA

Micheal L. Dent mdent@buffalo.edu

Department of Psychology
University at Buffalo
State University of New York (SUNY)
B76 Park Hall
Buffalo, New York 14260, USA

A Sound Plan for Attracting Students of Color

Tyrone Porter

The demographic survey completed by the Acoustical Society of America (ASA) in 2018 confirmed what many of us suspected, that the composition of the ASA membership does not reflect the demographics of the US population. This is particularly true with respect to Black representation because less than 2% of the membership that responded to the survey identified as Black.

The ASA Committee for Improving Racial Diversity and Inclusivity (CIRDI) that I chair was formed in the summer of 2020 (Porter, 2020, acousticstoday.org/porter-16-4) and charged with developing initiatives and activities to address this glaring problem within the Society and, most importantly, within academic programs and professions related to acoustics. One of the first questions CIRDI discussed was, “Why are there so few persons of color, particularly Blacks, in acoustics or acoustics-related fields?” Through our conversations, we recognized that there are few opportunities for Black students, especially undergraduate students, to be exposed to acoustics in a structured format. It is more likely that a Black student will discover acoustics and careers in the field through their own efforts rather than through a structured program (Scott, 2020, acousticstoday.org/ScottNewNormal). I share my own experience as an example of what the ASA must address to diversify the field and its membership.

I have been interested in physics and engineering since high school but was completely unaware of acoustics. Most of my high-school science classes focused on fundamentals (i.e., the biology of life across scales, Newton’s Laws), and I was only introduced to sound waves in my physics class. However, the introduction was very superficial, and the teacher never discussed careers in acoustics or acoustics-related fields.

On completing high school, I enrolled at Prairie View A&M University (PVAMU), which is a Historically Black College/University (HBCU) outside Houston, TX, and majored in electrical engineering. Similar to high school, the college physics course touched on acoustics and sound waves but with very little depth.

I was finally introduced to the fascinating world of acoustics during a summer research experience at Duke University (Durham, NC). The program was funded by the National Science Foundation, and I requested a research project in biomedical engineering to learn more about the field. Interestingly, my summer project focused on building and characterizing the performance of small transformers that would be installed in measurement devices for cardiac electrophysiology studies. I was and remain to this day an innately curious person, and so I would walk the hallways in the Pratt School of Engineering at Duke and read the research posters.

I discovered that the Duke Biomedical Engineering Program had a very strong diagnostic ultrasound group and found the research to be accessible for an electrical engineering student. On returning to PVAMU, I spent the year researching biomedical engineering graduate programs as well as companies that produced diagnostic ultrasound systems.

The following summer, I secured an internship in the Ultrasound Division of General Electric Medical Systems, which is now GE Healthcare. I had a very supportive supervisor and an extremely positive experience, which solidified my decision to pursue a career in biomedical ultrasound. My supervisor informed me of universities that had strong research programs in biomedical ultrasound, including the University of Washington (UW; Seattle). I was fortunate to be admitted to the bioengineering program at the UW, and I joined the research group led by Larry Crum. Larry recommended early in my graduate career that I join the ASA, and he served as a guide at its meetings. Larry also recommended that I attend programs that would provide additional instruction in acoustics while also expanding my network, such as the Physical Acoustics Summer School. I completed

https://doi.org/10.1121/AT.2021.17.1.65

my doctoral studies in 2003 and have been an active member of the ASA for more than 20 years.

It is important to note that my training in acoustics occurred predominantly after completing my undergraduate degree. However, if I did not take it on myself to learn about biomedical ultrasound as an undergraduate student, I never would have specialized in the area as a graduate student and I never would have joined the ASA.

Based on my experience, the CIRDI acknowledged that the ASA needs to create more opportunities for students of color to get introduced to acoustics and acoustics-related professions. The committee proposed that the ASA establish and manage a summer research and internship program in acoustics and acoustics-related fields for undergraduate students of color. In addition to funding the students, the ASA will provide a short course in acoustics in preparation of the summer experience.

Furthermore, ASA members will host virtual gatherings for the students to foster a community and discuss the academic and professional pathways available in acoustics and acoustics-related fields. The American Institute for Physics (AIP) awarded the ASA seed funding from its Diversity Action Fund to support launching the program in 2021. For many students, the summer program may be their first substantive experience with concepts, technologies, or processes involving acoustics. A positive experience both technically and culturally may serve as a first step toward pursuing a career in a field related to acoustics and becoming a member of the ASA. We are seeking mentors committed to diversifying their profession to host these aspiring young scholars. We plan to foster community among the mentors as well by hosting workshops and virtual gatherings to discuss and share best practices for mentoring students from underrepresented groups. Although not required, mentors and/or companies willing to fund a student will enable the ASA to include more students in the program. More information about the program and expectations for mentors can be found on the ASA Diversity Initiatives page (available at acousticalsociety.org/diversity-initiatives). If you are interested, please contact Tyrone Porter (tmp6@utexas.edu).

The summer program is part of a broader strategic plan crafted by the CIRDI to increase the representation of persons of color in acoustics-related careers and in the ASA. The plan includes various initiatives and activities designed to increase interaction and communication with students and professionals of color, such as working with faculty and administrators at minority-serving institutions to raise awareness of acoustics-related professions; organizing workshops and forums on topics related to diversity, equity, and inclusion; and creating a webpage to highlight persons of color in acoustics and acoustics-related fields. Please visit the ASA Diversity Initiatives page for more information and for opportunities to volunteer and implement the various initiatives.

References

Porter, T. (2020). Addressing the lack of Black members in the ASA. *Acoustics Today* 16(4), 75-76.

Scott, E. K. E. (2020). The need for a new normal. *Acoustics Today* 16(4), 77-79.

Contact Information

Tyrone Porter tmp6@utexas.edu

Department of Biomedical Engineering
The University of Texas at Austin
107 W. Dean Keeton Street
Austin, Texas 78715, USA

Hearing and Speech Research at the NIDCD

Debara L. Tucci

Introduction

From my position as director of the National Institute on Deafness and Other Communication Disorders (NIDCD) at the National Institutes of Health (NIH; Bethesda, MD), I am proud to lead an outstanding group of scientists and administrators who share my passion for scientific discovery and advancing public health in three program areas: hearing and balance; taste and smell; and voice, speech, and language. Our broad research portfolio of basic, translational, clinical, and public health research focuses on human communication and associated disorders.

At least 46 million people in the United States have a hearing or other communication disorder. I have dedicated my career to understanding the causes and impact of hearing loss and to developing treatments to restore hearing. Over my many years of clinical practice as an otolaryngologist surgeon-scientist, including more than 25 years at Duke University Medical Center (Durham, NC) where I cofounded the Duke Hearing Center and directed the medical center's cochlear implant program, I was privileged to care for and help many individuals with ear, hearing, and balance problems. I was also frustrated that our scientific understanding was insufficient to successfully treat every patient I encountered. As NIDCD director, it is gratifying to me to now guide the institute's exceptional biomedical workforce. I truly believe the research funded by our institute will continue to improve many lives in meaningful ways.

NIDCD: Three Decades of Discovery and Advancement

Over its 32-year history (available at bit.ly/3nGuuq3), NIDCD-supported researchers have made seminal advances in understanding the basic biology of sensory systems and disease mechanisms leading to increasingly effective, evidence-based treatments. Extraordinary research opportunities have led to scientific breakthroughs in the study of genes, proteins, cellular and molecular processes, neural circuits, and sensory and motor systems that directly affect our understanding of communication disorders. Current NIDCD-funded research promises to advance science in ways that directly impact patient care. Some examples include

- developing improved treatments for otitis media (middle ear infections);
- identifying and characterizing genes responsible for hereditary hearing impairment;
- studying genes associated with tumors affecting human communication;
- investigating gene therapy for treating hearing loss and dizziness;
- exploring the genetic bases of child language disorders as well as characterizing the linguistic and cognitive deficits in children and adults with language disorders;
- identifying biomedical and behavioral issues associated with communication impairment and disorders;
- researching improvements to assistive device technology that benefits those with hearing loss; and
- engineering a "thoughts into speech" algorithm for assistive communication devices to help people with amyotrophic lateral sclerosis, stroke, or neurodegenerative disease regain their ability to communicate.

NIDCD research has informed a practice that many now recognize as routine, universal newborn hearing screenings, and has supported research that helps us better understand the role taste and smell play in nutrition and health. Our national education campaign, "It's A Noisy Planet. Protect Their Hearing."® (available at noisyplanet.nidcd.nih.gov), has educated millions of teens, parents, and teachers about noise-induced hearing loss and how to prevent it. And our commitment to research that improves access to affordable hearing health care will help many Americans with hearing loss who could benefit from assistive hearing devices, such as hearing aids, but currently can't afford them.

NIDCD Budget and Spending Overview

For fiscal year (FY) 2020, Congress appropriated approximately $491 million to the NIDCD. This appropriation represents about 1.2% of NIH's total budget

https://doi.org/10.1121/AT.2021.17.1.67

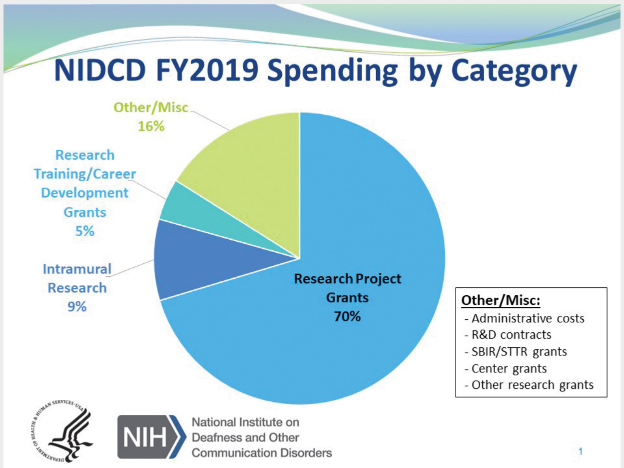


Figure 1. *This chart shows National Institute on Deafness and Other Communication Disorders (NIDCD) fiscal year (FY) 2019 spending by category: 70% research project grants; 16% other/miscellaneous, including administrative costs, research and development (R&D) contracts, Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) contracts, Center grants, and other research grants; 9% intramural research; and 5% research training/career development grants.*

($41.7 billion). NIDCD FY2020 appropriations are a 3.4% increase over the FY2019 budget.

Figure 1 shows an overview of FY2019 spending (the latest available) by category/funding mechanisms for intramural and extramural research programs. *Intramural* research is research conducted by scientists at NIH. *Extramural* research is research conducted by scientists at US and international research centers, universities, and medical centers.

Figure 2 notes the percentage of FY2019 intramural and extramural research spending for each of our seven mission areas.

A Broad Focus for NIDCD Funding Opportunities

The NIDCD distributes its resources among many diverse programs and mechanisms. The institute is committed to funding the largest number of meritorious projects possible while allowing the flexibility needed to support selected program priorities and respond to emerging scientific opportunities. Our funding applicants and grantees represent diverse professional and academic programs, from biology and medicine to engineering, physics, and mathematics — that address research questions relevant to the NIDCD's multidisciplinary mission.

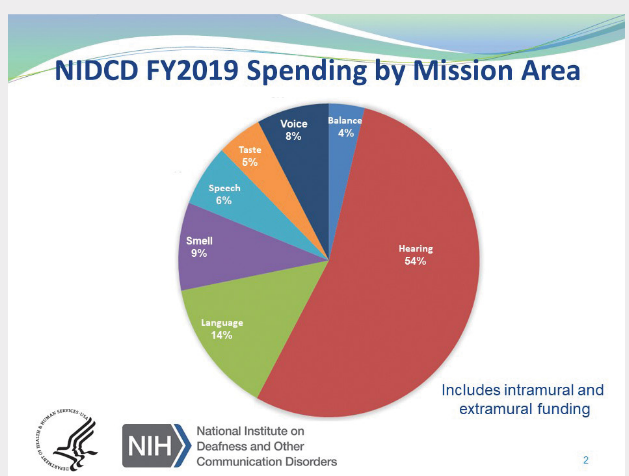
Examples of NIDCD support that may interest applicants from traditional and nontraditional biomedical fields include the NIDCD Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) programs. These funding mechanisms aim to stimulate technology and research innovation with a potential for product commercialization. VocaliD is a personalized-voice product designed by NIDCD voice scientists and is now available to people with severely impaired speech. The device blends the speech of two individuals, a donor and a recipient, to recreate the recipient's natural voice.

The NIDCD establishes general guidelines for funding based on scientific merit, responsiveness to the institute's priorities, and availability of funds. I encourage individuals from nontraditional and traditional biomedical areas connected to our program areas to consider NIDCD as a source for research support. Please see our guidelines for more information about our funding opportunities (available at www.nidcd.nih.gov/funding).

Training the Next Generation for Success

We continue the NIDCD's long commitment to a research environment that supports scientists in a variety of related disciplines and at all stages of their professional

Figure 2. *This chart represents the percentage of FY2019 funding across the 7 mission areas of the NIDCD: 54% for hearing, 4% for balance, 5% for taste, 9% for smell, 8% for voice, 6% for speech, and 14% for language.*



careers. My hope is that trainees funded by the NIDCD will subsequently submit successful career development applications and continue a trajectory to productive and fulfilling research careers. As an otolaryngologist surgeon-scientist, I am committed to training the next generation of otolaryngologist researchers (available at bit.ly/NIDCD_OSSP) who can leverage their unique clinical experience and research skills to address important questions in human disease and disorders.

The NIDCD supports a variety of grant mechanisms that are tailored to support different stages of professional career development. Support for investigators who have received their terminal education degrees within the past 10 years (early-stage investigators; ESIs) is reflective of our commitment to early career development. The institute has a long history of supporting ESIs through special programs, including the Early Career Research Award (ECR R21), and an expedited review of predoctoral and postdoctoral fellowship applications. Our training programs are designed to support the next generation of scientists and other professionals who will address tomorrow's expanding health care needs. I encourage you to peruse our extensive research training and career development opportunities at the NIDCD website (available at www.nidcd.nih.gov/training).

Commitment of the NIDCD to a Diverse Biomedical Workforce

The NIDCD has diligently worked to increase the diversity of the research pipeline across our mission areas. When scientists and trainees from different backgrounds work together, their unique perspectives and experiences stimulate creativity and innovation, yielding higher quality research than less diverse teams (available at bit.ly/3nJ4zhC). Importantly, diverse research teams are more likely to ensure that members of underserved populations will support and participate in research and that the research we invest in addresses questions that are meaningful to these communities. Increasing scientists' understanding of disparate groups benefits us all and is at the core of the NIH mission: *to uncover new knowledge that will lead to better health for everyone*. Deafness and other communication disorders, after all, cross all cultural, racial, and gender boundaries. Despite these efforts, however, the proportion of investigators receiving funding in our mission areas who are members of underrepresented minority groups remains small.

To affirm the NIDCD's commitment to inclusive excellence and our resolve to both embrace and enable the contributions of a diverse scientific workforce, I initiated several steps to ensure that our commitment has an impact. Together with our scientific advisory council and other stakeholders, the NIDCD is exploring how we can most effectively engage underrepresented minority scientists throughout their careers and support training, mentoring, and leadership development programs to ensure a robust and diverse workforce. Furthermore, we are looking at how best to increase participation of underrepresented minority populations in research studies in our mission areas.

Supporting Research Toward Affordable, Accessible Hearing Health Care and Improving Global Hearing Health

Approximately 15% of US adults report some degree of hearing loss. Untreated hearing loss is a significant public health issue. Higher total health care costs, a higher risk of dementia and cognitive decline, falls, depression, and a lower quality of life have been associated with untreated hearing loss in older adults (Deal et al., 2019). As the lead federal agency supporting research to prevent, detect, and treat hearing loss, the NIDCD supports initiatives to improve access to affordable hearing health care (available at bit.ly/330QIeE). One example is NIDCD's contributions to and major support for the National Academies of Sciences, Engineering, and Medicine consensus study, "Hearing Health Care for Adults: Priorities for Improving Access and Affordability" (2016). Cosponsored by the NIH through the NIDCD and the National Institute on Aging, as well as four other federal agencies and a nonprofit patient advocacy group, the study concluded that the diverse needs of adults with hearing loss were not being met. As a result, one of the independent panel's 12 recommendations for improving adult hearing health care was for the Food and Drug Administration (FDA) to create and regulate a new category of over-the-counter (OTC) hearing devices for adults with mild-to-moderate hearing loss. These products are expected to come to market soon, pending release by the FDA of the final regulations for guidelines and quality standards. Additionally, a small-business research grant from the NIDCD led to the first self-fitting hearing aid approved by the FDA in 2018. The NIDCD remains committed to improving the landscape of adult hearing health care and encourages continued research to fill remaining gaps.

Global hearing health care is another NIDCD priority and one that also embraces multidisciplinary approaches. I cochair The Lancet Commission on Hearing Loss (available at globalhearinglosscommission.com), which pursues innovative ideas that challenge the accepted thinking on identification and treatment of hearing loss worldwide. The commission seeks to develop creative approaches focused on policy solutions and the use of new technologies and programs to enable those with hearing loss worldwide to be fully integrated into society. We will share our findings in spring 2022. I encourage you to learn more about the NIDCD's commitment to global health (available at bit.ly/3kMEhZL).

NIDCD's Strategic Plan for 2022–2026

One current focus is the development of the institute's strategic plan for 2022–2026. Throughout this process, we are formulating ambitious yet achievable goals for research in our mission areas, goals that will further our scientific understanding of basic biological systems, human disease mechanisms, and promising treatments. We will continue to prioritize accessibility of care and research dissemination as core components of our mission. We are committed to making treatments accessible to all, using a full range of innovative technologies and approaches to help all populations, inclusive of gender, race, ethnicity, socioeconomic status, geographic location, and communication method. We also value our work with the many individuals and groups outside of the NIH who represent those affected by deafness and other communication disorders.

Looking to the Future

I am optimistic that in the coming years we will make tremendous progress in addressing the scientific and clinical challenges related to the mission areas of the NIDCD. I am especially proud of how our workforce successfully navigated pandemic-imposed challenges with dedication and spirit during my first year as NIDCD director. I am excited about continuing this work and commitment with my colleagues in the months and years ahead to apply this energy across our mission areas. Together, we will continue to see the quality of life improved by our research.

References

Deal, J. A., Reed, N. S., Kravetz, A. D., Weinreich, H., Yeh, C., Lin, F. R., and Altan, A. (2019). Incident hearing loss and comorbidity: A longitudinal administrative claims study. *JAMA Otolaryngology-Head & Neck Surgery* 145(1), 36-43. Available at https://bit.ly/3325NfZ.

National Academies of Sciences, Engineering, and Medicine (2016). *Hearing Health Care for Adults: Priorities for Improving Access and Affordability*. The National Academies Press. Washington, DC. Available at https://bit.ly/3pGJbeu.

Contact Information

Debara L. Tucci debara.tucci@nih.gov

Office of the Director
National Institute on Deafness and Other Communication Disorders (NIDCD)
Building 31, Room 3C02
31 Center Drive, MSC 2320
Bethesda, Maryland 20814, USA

Obituary

James David Miller, 1930–2020

James David (Jim) Miller, a Fellow and member of the Executive Council of the Acoustical Society of America and prolific contributor of research in several areas of human communication, died at age 90 on August 20, 2020, in Bloomington, IN. Jim's career spanned a remarkable 65 years; his final publications in *The Journal of the Acoustical Society of America* in 2017 and 2020 are extensive studies of speech recognition by hearing aid users.

Born in West Allis, WI, Jim attended the University of Wisconsin-Madison, where he was an assistant in Harry Harlow's primate laboratory. He went on to graduate work at Indiana University (Bloomington), where his dissertation on noise-induced temporary threshold shifts was supervised by James Egan. In 1958, after completing his PhD in experimental psychology, he was placed in charge of an Air Force-funded project on noise-induced deafness in cats. After three years, that effort resulted in one of the most widely cited studies of the systematic influence of various levels of intense noise on the mammalian ear, in terms of both behaviorally measured hearing loss and damage to specific cochlear structures.

In 1961, Jim was recruited by the Central Institute for the Deaf (CID) in St. Louis (MO). He stayed at the CID for the next 40 years, initially as head of the animal research program and later assuming many other leadership roles. He retired in 2001 as Emeritus Director of Research at the CID. In 2003, he was invited for a semester visit to Indiana University and stayed in Bloomington for the remaining 17 years of his life, as an adjunct faculty member and as Principal Scientist at Communication Disorders Technology, Inc.

The list of Jim's scientific accomplishments is long, but even longer is that of the many students and research collaborators who have been outspoken in describing his excellence as a teacher, his remarkable insights in the interpretation of research findings, and his steadfast adherence to rigor and honesty in science. He pioneered the use of the chinchilla in auditory research and fostered the work of Robert Dooling on songbirds and Patricia Kuhl on sensitivity to speech sounds in infants. With anatomists Walter Covell and Barbara Bohne, he extended his groundbreaking work on noise-induced deafness in animal models. As director of research, he fostered and guided a group of CID audiologists and engineers who patented the first wearable digital hearing aid. Jim also became interested in speech recognition and, with various collaborators, published a series of studies, gradually developing what became his "auditory-perceptual theory of phonetic recognition." In Jim's final years of research, he directed a multi-university study of speech perception training by hearing aid users plus an application of the same principles to improving the communication skills of students of foreign languages.

In addition to his remarkably full career as a scientist, Jim found time for a lifelong dedication to becoming the best tennis player he could. He also firmly believed in "striving for a more perfect union" and did house-to-house campaigning in recent elections.

He is survived by his former wife Dolores; their children, Valerie, Lucia, and Harry; and granddaughter Rose.

Selected Publications by James David Miller

Miller, J. D. (1989). Auditory-perceptual interpretation of the vowel. *The Journal of the Acoustical Society of America* 85, 2114-2134.

Miller, J. D., Niemoeller, A. F., Pascoe, D., and Skinner, M. W. (1980). Integration of the electroacoustic description of hearing aids with the audiologic description of clients. In G. A. Studebaker and I. Hochberg (Eds.), *Acoustic Factors Affecting Hearing Aid Performance*. University Park Press, Baltimore, MD.

Miller, J. D., Watson, C. S., and Covell, W. P. (1963). Deafening effects of noise on the cat. *Acta Oto-Laryngologica Supplement* 176, 1-91.

Miller, J. D., Watson, C. S., Leek, M. R., Dubno, J. R., Wark, D. J., Souza, P. E., Gordon-Salant, S., and Ahlstrom, J. B. (2017). Syllable-constituent perception by hearing-aid users: Common factors in quiet and noise. *The Journal of the Acoustical Society of America*. 141(4), 2933-2946.

Miller, J. D., Watson, C. S., Leek, M. R., Wark, D. J., Souza, P. E., Gordon-Salant, S., Ahlstrom, J. B., and Dubno, J. R. (2020). Sentence perception in noise by hearing-aid users predicted by syllable-constituent perception and the use of context. *The Journal of the Acoustical Society of America*. 147, 273-284.

Written by

Charles S. Watson watson@indiana.edu

Indiana University, Bloomington

Obituary
Jeffrey A. Nystuen, 1957–2020

Jeffrey A. Nystuen, an admired colleague in the acoustics and oceanography communities, recently passed away. Jeff was a Principal Oceanographer at the Applied Physics Laboratory (APL) at the University of Washington (UW), Seattle, from 1995 until his retirement in 2018. Jeff was the recipient of the 2003 Medwin Prize of the Acoustical Society of America (ASA) and was named ASA Fellow in the same year. Jeff was born in Seattle, grew up in Ann Arbor, MI, and was a graduate of the University of Michigan and Scripps Institute of Oceanography/University of California, San Diego, La Jolla.

Jeff is widely regarded as a pioneer in acoustic rainfall measurement. By monitoring the sound of rain falling on the ocean, Jeff was able to learn about patterns of rainfall and cloud formation over the open ocean. Knowing more about the role of rainfall has given meteorologists a better understanding of weather phenomena such as El Nino and the floods and droughts that it triggers around the world.

Jeff's interest in acoustic rain measurement was inspired by the suggestion from his doctoral advisor Robert Stewart and coadvisor Walter Munk at Scripps. The remote sensing of satellite/radar rain measurement provides large surface coverage, yet the spatial resolution is poor and needs local ground truth for the calibration. To measure the "local" rain rate at sea is no easy task due to the destructive force at the air-sea interface. The passive acoustic method provides an alternative to measuring the rainfall away from the air-sea interface and with a much larger surface coverage area than the conventional accumulation-type rain gauge on a surface buoy. Because the rain pattern is intermittent and varies spatially, the passive acoustic method is better, in theory.

While as a faculty member at the Naval Postgraduate School in Monterey, CA, Jeff studied the sound of individual raindrop splashes with Herman Medwin. They used an abandoned elevator shaft as an acoustic chamber to drip various sizes of water drops and recorded the sound generated individually. They discovered a new sound-generating mechanism due to the bubble entrapment and improved predictions of the underwater sound produced.

In 1995, Jeff moved to the APL and later became an affiliate faculty member at the School of Oceanography at the UW. There he developed a passive acoustic recording system called Passive Aquatic Listeners (PALs). He advocated for a smaller data size through intelligent sampling schemes. He established a field program to measure rainfall at sea using his acoustical technique and collaborated with colleagues around the world.

Jeff was a generous and insightful colleague, collaborator, advisor, and mentor. His perspective gave the sound of rainfall a special meaning. Whenever we hear raindrops on water, we think of him.

He is survived by his parents in Ann Arbor, sister and brother-in-law in Massachusetts, and many friends in Seattle and around the world.

Selected Publications by Jeffrey A. Nystuen

Ma, B. B., Nystuen, J. A., and Lien, R.-C. (2005). Prediction of underwater sound levels from rain and wind. *The Journal of the Acoustical Society of America* 117, 3555-3565.

Nystuen, J. A. (1986). Rainfall measurements using underwater ambient noise. *The Journal of the Acoustical Society of America* 79, 972-982.

Nystuen, J. A. (2001). Listening to raindrops from underwater: An acoustic disdrometer. *Journal of Atmospheric and Oceanic Technology* 18, 1640-1657.

Nystuen, J. A., and Medwin, H. (1995). Underwater sound produced by rainfall: Secondary splashes of aerosols. *The Journal of the Acoustical Society of America* 97, 1606-1613.

Written by:

Barry Ma barry@apl.uw.edu
Applied Physics Laboratory and School of Oceanography
University of Washington, Seattle

Peter Leopold peterleopold23@gmail.com
BioAnalyte Inc.
Portland, ME

Obituary
Ann Kristen Syrdal, 1945–2020

Ann Kristen Syrdal passed away on July 24, 2020, at her home in San Jose, CA. In a career spanning five decades, Ann made outstanding contributions to our understanding of human speech perception and to the development of natural-sounding text-to-speech (TTS) synthesis.

Ann received her PhD in psychology from the University of Minnesota (Minneapolis) in 1973, concentrating on human speech perception. Categorical perception, the way human listeners partition a continuum of varying acoustic patterns into discrete linguistic categories, drove Ann's early work as did the *invariance problem*: how vastly different acoustic patterns such as speech uttered by a young child and an adult male can be mapped to the same linguistic elements. Working with colleagues at Haskins Laboratories and in her dissertation, Ann explored categorical perception for stop consonants using speech synthesis to generate stimuli, contributing to our understanding of phoneme recognition. Additionally, Ann collaborated with then-husband Robert Lasky, demonstrating that universal categorical perception in young infants must later be "tuned" to the specifics of the infant's native language.

One of Ann's most important contributions, supported by a National Institutes of Health Career Development Award, addressed the invariance problem and speaker-independent classification of vowels. Syrdal and Gopal presented a model in which differences measured on a Bark scale between a speaker's fundamental frequency and pairs of formants leads to the accurate classification of vowels in the front-back and high-low dimensions used to describe vowels in articulatory phonetic terms.

Ann moved to AT&T Bell Labs in Naperville, IL, where she worked on aspects of speech synthesis, including duration modeling and, most importantly, developing the first female synthetic voice, work for which she was eventually honored by being made an Acoustical Society of America (ASA) Fellow. For her applied research, she was honored as a Distinguished Member of the Technical Staff. This work also informed a volume on applied speech technology that sampled most evolving applications of speech technology and reflected her particular interest in the use of speech technology in assistive communication aids for persons with sensory or motor deficits that present challenges to auditory or oral communication.

Later, Ann joined the AT&T Next-Gen TTS project, developing state-of-the-art natural-sounding corpus-based speech synthesis. Ann guided the preparation of large speech corpora, designed voices, and carried out detailed experimental evaluations that led to higher quality. The results of the work included new high-quality female voices and an African American TTS voice. The system was commercialized and enjoyed lasting success under the product name AT&T Natural Voices. Ann also led the ASA committee that developed the American National Standards Institute (ANSI)/ASA standard on evaluating the intelligibility of TTS systems. Ann was a wonderful friend to many AT&T colleagues and mentor to young researchers.

Selected Publications by Ann Kristen Syrdal

Beutnagel, M., Conkie, A., Schroeter, J., Stylianou, Y., and Syrdal, A. (1999). The AT&T Next-Gen TTS system. In *Proceedings of the Joint meeting of the Acoustical Society of America (ASA), European Acoustics Association (EAA), and German Acoustics DAGA Conference*, Berlin, Germany, March 14-19, 1999, pp. 18-24. https://acousticstoday.org/SyrdalATT.

Lasky, R. E., Syrdal-Lasky, A., and Klein, R. E. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology* 20(2), 215-225.

Mattingly, I. G., Liberman, A. M., Syrdal, A. K., and Halwes, T. (1971). Discrimination in speech and nonspeech modes. *Cognitive Psychology* 2(2), 131-157.

Syrdal, A. K., Bennett, R. W., and Greenspan, S. L. (Eds.) (1994). *Applied Speech Technology*. CRC Press, Boca Raton, FL.

Syrdal, A. K., and Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America* 79(4), 1086-1100.

Written by

H. Timothy Bunnell tbunnell@nemours.org
Center for Pediatric Auditory and Speech Sciences
Nemours/Alfred I. duPont Hospital for Children
1600 Rockland Road, Wilmington, DE

Alistair D. Conkie aconkie@apple.com
Apple, Inc., Cupertino, CA

Advertisers Index

Advertising Sales & Production

Debbie Bott, Advertising Sales Manager
Acoustics Today, c/o AIPP, Advertising Dept
1305 Walt Whitman Rd, Suite 300, Melville, NY 11747-4300
Phone: (800) 247-2242 or (516) 576-2430
Fax: (516) 576-2481 | Email: dbott@aip.org

Business Directory

For information on rates and specifications, including display, business card and classified advertising, go to Acoustics Today Media Kit online at: https://publishing.aip.org/acousticstodayratecard *or contact the Advertising staff.*

TYPE 4966 MICROPHONE FAMILY

HIGH-PRECISION
1/2'' FREE-FIELD MICROPHONE
HANDMADE IN DENMARK

To help engineers and acousticians achieve accurate measurements in their everyday tasks, Bruel & Kjær has introduced the Type 4966 family of robust, reliable, free-field microphones ideal for high-precision acoustic measurements.

The Type 4966 family has 4 microphones for different applications:
1. For all-round acoustic measurements in temperatures up to 125°C
2. Well suited when low-frequency disturbance, such as body boom or road noise, needs to be eliminated
3. For measurements in confined spaces or when building an array with a low noise floor
4. For high sound pressure level (SPL) measurements or if the integrity of the microphone measurement channel needs to be checked

Bruel & Kjaer North America Inc.
3079 Premiere Parkway, Suite 120 Duluth, GA 30097
Telephone: +1 770-209-6947
Email: bkinfo@bksv.com

Brüel & Kjær
an HBK company

www.bksv.com/4966-family