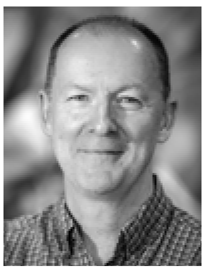
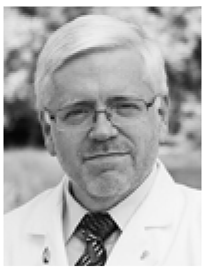
Scott McCoy, Associate Editor

Practical Science in the Studio, Part 2: "Low-Tech" Strategies

John Nix, David Meyer, Ron Scherer, and Deirdre Michael



John Nix



David Meyer

Ron Scherer

Deirdre Michael

Journal of Singing, March/April 2021
Volume 77, No. 4, pp. 509–513

[Editor's Note: This article, an initiative of the NATS Voice Science Advisory Committee, is the second in a three-part series that seeks to introduce the reader to practical and cost-effective strategies for using science to enhance singing instruction. The three articles in this series are divided into "no-tech," "low-tech," and "high-tech" segments.]

As of October 2020, singers and teachers around the world have faced nine months of artistic, pedagogic, and economic challenges brought on by the global Coronavirus pandemic. New ways of performing, teaching, and earning a living had to be adopted on short notice. Technological means to continue to work while following health guidelines became essential for the singing profession. Routers, Internet speeds, 100 foot ethernet cables, low latency peer-to-peer connections, HEPA filters, and N95 masks became common topics of conversation in NATS social media discussions. For some teachers, this forced, lightning fast transition to a more technologically driven means of teaching was very daunting.

It seems timely, then, that the NATS Voice Science Advisory Committee launch a three-part series of articles ("no-tech," "low-tech," and "high-tech") on the practical use of technology in voice pedagogy during this period.[1] This article, the second in the series, considers "low-tech" solutions, such as the nuts and bolts of digital audio and video recording, cellphone apps, and freely available web resources to improve perceptual skills. By low-tech we mean tools that most teachers already have in their studios and may be quite familiar with using.

BASICS OF RECORDING

Making quality audio and video recordings has become an essential part of the lives of singers and teachers. We make recordings for asynchronous instruction, practice monitoring, lesson documentation, online auditions, and prerecorded performances. A good recording can be the difference between a student understanding or not understanding a subtle teaching concept or between getting and not getting a gig or an apprenticeship. What are some simple steps to improve the quality of recordings?

A few quick definitions are needed before we delve into this topic. When we make recordings, we convert a continuous analog signal (fluctuations in an electric current from the microphone) into a digital signal (strings of 0s and 1s) that a computer can process. Computers have sound cards that perform this analog to digital conversion process. How many times per second the computer measures the incoming analog signal and converts it into a digital signal is called the sampling rate. Think of it like taking pictures of a growing teenager: If you took only one picture per year, you would see a very abrupt change in the young person's height and body dimensions from picture to picture; if you took a photograph every day, you would see changes happening very gradually and smoothly. So a high sampling rate enhances the quality of the recording, particularly higher frequency sounds, but it also makes the recorded files much larger. Most computers default to 44,100 samples per second, as you want a sampling rate at least double the highest possible sound in the recording, and human hearing tends to end at 22,000 Hz. You can hear the difference in successively higher sampling rates by accessing this video: https://youtu.be/4j3miQZpVu4.

Another term often seen with computers and sound processing is "bit depth." With recordings, this refers to how many levels or values of amplitude can be represented. The greater the bit depth is, the better the resolution of the variations in intensity. Most computers will record at 16 bit resolution, giving an amazing 65,536 different levels possible (2 to the 16th power).[2] To follow the analogy of measuring the growing teenager, if your measuring tape had markers only every meter, you could only approximate how tall the youngster was, while if your measuring tape had millimeter lines, you could track his or her growth rate very precisely.

Thus far, it is clear that to make good recordings you want a high sampling rate and at least 16 bit resolution. What kind of file should you use when you make a recording? For the original master recording, you should use a lossless file format, such as .wav, FLAC, ALAC, or APE. By lossless, we mean that none of the data are thrown out when the signal is saved. Once the data are lost in a conversion process to another file format, they can never be recovered, so master recordings should be made with a high sampling rate, at least 16 bit resolution, and in a lossless format. The drawback of lossless format .wav files is that they are large files and as such take a longer time to download: a one minute stereo .wav recording at 44,100 samples per second and 16 bit resolution is 10.6 MB (megabytes) in size.[3] So while master recordings should be .wav files, files for usage on a website (where fast loading is a priority) should be in one of the compressed formats, such as .mp3, .aac, or .wma. The compressed formats are smaller in size, so they load fast in web browsers, but the compression process discards a lot of the original data. Most computers allow you to select the bit rate when you convert to a compressed format like .mp3; higher bit rates provide higher sound quality, as less of the original data are discarded in compressing the file. File sizes for a one minute .mp3 file typically range from 1.0–2.4 MB.[4] What about video? Common file types are .mov (Apple devices), .wmv (Windows devices), and .mp4, all of which are compressed to some degree to save space and speed loading time. Of the three, .mp4 plays on the widest variety of devices. Typically, the type of camera used determines with which file type you record.

Microphones are the most important part of your "signal chain" from performer to stored recording. Microphones can be described by their type, sensitivity pattern, and frequency response. The two primary microphone types are dynamic and condenser microphones. Dynamic microphones are cheaper to purchase, more durable, and less sensitive; they are great for public speaking and for use as handheld microphones on stage singing in contemporary styles. Condenser microphones are more expensive to purchase, more fragile, and much more sensitive; as such, they are great for studio recordings and research use. Microphone patterns are either directional (the microphone is sensitive in certain directions but not others) or omnidirectional. With regard to frequency response, some microphones boost the intensity of certain frequencies, while others, so-called flat response microphones, avoid this boosting factor. For more information about microphones, readers may find Brian Manternach's interview with NATS Voice Science Advisory Committee member Ian Howell helpful (https://www.csmusic.net/content/articles/mic-check/).

One final piece of low-tech equipment is the cable used to connect the microphone to the computer. The two most common types in use are XLR and USB cables. XLR cables are "quieter," that is, they carry nothing but the audio signal, but they require an audio interface of some sort to connect them with a computer. What is an audio interface? Audio interfaces are small box-like devices that have XLR and ¼ inch phono microphone or line inputs, gain controls for each input channel, a headphone jack and headphone volume control, an analog to digital converter, and a variety of outputs, including some type of USB output that can be connected to a laptop or desktop computer. USB cables are fairly ubiquitous and provide easy connection with computers, but they have a higher noise "floor" due to powering the recording device as well as sending the audio signal.

To sum up: For studio quality recording, use a high sampling rate (at least 44,100 samples per second), at least 16 bit resolution, .wav or other lossless file types, a condenser microphone with a flat response, an XLR cable, and an external audio device (common brands include Presonus, Audient, and Focusrite). For putting files on the internet for rapid loading and playback, consider a compressed audio file format like .mp3, .acc, or .wma, and a widely used video file type like .mp4.

CELL PHONE SOUND LEVEL METER APPS

Smartphone apps can be useful in teaching voice, particularly when distance instruction is involved. Many of us have been teaching via Zoom, Facetime, Skype, and similar apps, and certainly smartphones have been an integral part of our new pedagogic practice. One application that teachers may not be aware of is the monitoring of how loudly our students are singing.

Simply said, loudness matters. Oversinging (vocal hyperfunction) can increase the risk of vocal injury, and undersinging (vocal hypofunction) can prevent students from developing the stamina to meet the stylistic and/or marketplace demands of their chosen repertoire. Functional, optimal training requires finding that "sweet spot" where singers are fully exercising their instrument, but avoiding hyperfunctional vocal strain.

In voice science the acoustic power of a sound is related to its *intensity*. This intensity is measured in decibels (dB) and is what we perceive as the *loudness* of sounds. Because it is a perceptual construct, vocal loudness can be difficult for teachers to assess. When lessons are not in-person, the singer's audio equipment will alter what the teacher hears, often limiting or distorting loud sounds. Even when lessons are in-person, perceptions of vocal loudness may differ between the teacher and student. Sound pressure level (SPL) smartphone apps can help to reduce this perceptual subjectivity.

Finding and using an SPL app can be a relatively simple task: a quick internet search (suggested keywords: SPL, decibel, dB, meter) will result in many apps of varying complexity and cost. Decibel-X and SPL meter are widely used, but there are many others. To use an SPL app in voice lessons, the teacher needs to do three things: 1) teach the student to maintain a constant distance from the phone while singing; (2) teach the student to use the same app with the same settings in the same acoustic environment for consistent measurements; and (3) set anchors for the student's ideal vocal loudness.

Maintaining a constant distance between singer and smartphone is important because intensity measures will change depending on how far the microphone is from the sound source. Double the distance between the singer and the phone, and the measurement will drop by approximately 6 dB; reduce the distance between the singer and the phone by one half and the measurement will increase around 6 dB. An easy way to ensure that this distance stays constant is to have students place their smartphone on a music stand or on their laptop propped against the screen. Then students should extend their arm and stand so that their fingers can just barely touch the phone. Singing at an arm's length from the smartphone is a simple, repeatable method for maintaining a constant distance.

Next, when using a sound level meter app in lessons, find an app that works well and use it consistently, otherwise the reliability of the measurements will suffer. Work with students to find an app and settings on that app that provide consistent results, then stick with them! Furthermore, if you are using an app, be sure that students sing in the same room in the same place every time. Sound level readings are very environment-sensitive (just compare your own singing in your bathroom versus your bedroom), so the more constant the conditions, the more reliable and meaningful the readings will be for your students' singing and your teaching.

Loudness anchors give students an idea of how loudly they *should* be singing. These anchors are most easily set during in-person lessons, but distance learners can be creatively accommodated. A simple vocal exercise (e.g., five note ascending and descending scales) in the middle of the singer's range can be practiced *a cappella* with the teacher while using the app. It is important to remember that accompaniments will add to the exercise's measured loudness, and may mask the singer's actual loudness. The decibel measure will bounce up and down, depending on the intensity of the singer's voice. Consider using an app with an analog meter, as a needle moving up and down is visually simpler and less potentially distracting than apps with flashy numerical displays. While the singer is practicing, the teacher could say, "When you are singing this pattern, try to keep the dB reading between 80 and 90" (or whatever range the teacher considers ideal for the student's vocal needs). Trial and error will allow teacher and singer to come to a consensus of the optimal range of dB readings for that singer, app, and acoustic environment. It is important to note that there is no *absolute* dB value that is correct. Sound pressure level (SPL) varies greatly across different contexts, and research-grade measurements are not necessary. Rather, consistency of measurement is what provides value to the student. A switch from a dorm room to a practice room could change the reading, confusing the student. Of course, the student could determine the optimal dB reading in each of a variety of practice settings, which may help develop reliability of loudness judgment for the student.

DEALING WITH VOICE QUALITY DEVIATIONS

Some singing students come to the singing teacher's studio with an exceptionally clear speaking and singing voice, and some with a distinctly unclear voice. It is difficult to determine if an unclear voice is due to a functional vocal technique issue that can be helped with singing lessons, or if the unclear voice is due to a vocal fold tissue pathology. It is ethically and physiologically important to refer a student with a deviant voice quality to a laryngologist before giving singing lessons. The singing teacher needs to know if there is a vocal fold tissue problem or other anomaly that requires medical attention before giving singing lessons.

Voice quality deviations have a variety of descriptors. The singing student with a deviant voice quality may have a sound that is rough, breathy, or strained, or may have a deviance in pitch or loudness, or may use excessive vocal fry, sound gurgly, etc. Following the Consensus Auditory-Perceptual Evaluation of Voice[5] (CAPE-V, ASHA Special Interest Group 3, Voice and Voice Disorders),

- a rough voice is a "perceived irregularity in the voicing source [larynx]";
- a breathy voice is when you hear an "audible air escape in the voice";
- a strained voice is the "perception of excessive vocal effort (hyperfunction)."

These three descriptions are based on changes that ostensibly occur at the level of the vocal folds, and are different from the qualities that are based on characteristics of the vocal tract, such as "backward" or "hypernasal." These basic quality deviations may be quite minor in many students, for which singing lessons improve the voice. But when these qualities, or pitch or loudness deviances, are relatively excessive, the referral to a laryngologist is important.

A minor problem with breathiness may be solved quickly with training a little more adduction in the singing voice, but if the change is not easy, referral again is important; there could be a vocal fold lesion (nodules, polyp, other) or neuromuscular weakness (paresis). Roughness also may pass with training toward a clear voice, but again if not soon resolved, the referral is important because roughness results from the inability of the vocal folds to function normally. Strain in the voice may be reduced with relaxation and better coordination of the muscles and configuration of the larynx with singing lessons, but if strain persists, there could be more involved than simple muscle tension, and referral again is important. Note that any voice condition needing medical attention should lead to a consultation with a voice team, which consists of a laryngologist (for medical diagnosis), a speech-language pathologist who specializes in voice disorders (for therapy), and the singing teacher (for singing lessons), among other professionals.

From a "low-tech" point of view, the ear of the singing teacher is the measuring tool. Do singing teachers need to train themselves to be able to rate the voice quality of singing students the same way as speech-language pathology students train to diagnostically assess the speaking voice? If the singing teacher would like to experience the training of a speech-language pathology student, the teacher may wish to follow the sequence offered by the University of Wisconsin-Madison, Voice Disorders: Simulations (https://slpsims.csd.wisc.edu/simulations.html), where there are many voice cases to listen to and rate for their deviancy in numerous categories. Another training opportunity for the CAPE-V assessment has been made available by the Voice Foundation. It can be found at https://voicefoundation.org/health-science/videos-education/pvqd/.

FINAL THOUGHTS

We hope these ideas will help singing teachers incorporate more technology into online, studio, and classroom teaching. In the next issue of the *Journal of Singing*, "high-tech" tools such as spectrographic software, the Madde voice simulator, and several others will be discussed in this column.

NOTES

1. David Meyer and Lynn Helding, "Practical Science in the Studio: "No-Tech" Strategies," *Journal of Singing* Volume 77, no. 3 (January/February 2021): 359–367.
2. "Implications of number of bits"; http://hyperphysics.phy-astr.gsu.edu/hbase/Audio/digit.html (accessed October 29, 2020).
3. The Audio Archive, "State–of-the-Art-Audio Transfer: audio bit rate and file size calculators"; http://www.theaudioarchive.com/TAA_Resources_File_Size.htm#File_Size_Tutorial (accessed October 29, 2020).
4. AudioMountain.com, "Audio file size calculations"; http://www.audiomountain.com/tech/audio-file-size.html (accessed October 29, 2020).
5. Gail B. Kempster, Bruce R. Gerratt, Katherine Verdolini Abbott, Julie Barkmeier-Kramer, and Robert Hillman, "Consensus Auditory-Perceptual Evaluation of Voice: Development of a Standardized Clinical Protocol," *American Journal of Speech-Language Pathology* 18, no. 2 (May 2009): 124–132.

John Nix, tenor, is Professor of Voice and Pedagogy at UTSA. He holds degrees from the University of Georgia, Florida State University, the University of Colorado, and a Vocology Certificate from the University of Iowa. Mentors include Barbara Doscher and Ingo Titze. His students have sung with Santa Fe Opera, Opera Theatre of Saint Louis, Chautauqua Opera, Opera Omaha, Arizona Opera, Nevada Opera, San Antonio Opera, and the Metropolitan Opera Chorus. He has won grants from The Grammy Foundation and the National Institutes of Health, the 2006 Van Lawrence Award, published more than 45 articles, and edited or contributed to 5 books, including *The Oxford Handbook of Music Education* and *The Oxford Handbook of Singing*.

A leading scholar and researcher of the singing voice, **Dr. David Meyer** is an active performer, teacher, clinician, and voice scientist. He serves as associate professor of voice and voice pedagogy at Shenandoah Conservatory, and is Director of the Janette Ogg Voice Research Center. He is also a member of the Scientific Advisory Board of the Voice Foundation and is co-chairman of the Voice Science Advisory Committee of the National Association of Teachers of Singing. In 2010 he received the Van L. Lawrence Fellowship, a prestigious national award in recognition of his contributions to the field of teaching singing and the use of voice science. Dr. Meyer's students have won numerous awards and have sung in major venues worldwide.

Ron Scherer, PhD, Distinguished Research Professor in the Department of Communication Sciences and Disorders, Bowling Green State University, teaches voice disorders and voice and speech science courses. His research interests include the physiology, mechanics, and acoustics of basic, abnormal, and performance sound production, and physical and computer modeling of phonation. He is a Fellow of the Acoustical Society of America, the American Speech Language Hearing Association, and the International Society of Phonetic Sciences.

Deirdre D. ("D.D.") Michael has been a singer all her life, a singing teacher for over 40 years, and a speech language pathologist since 1991. She has a BA in music, and an MA in speech-language pathology and PhD in communication disorders, specializing in voice science, both from the University of Minnesota. She is an Assistant Professor in the Department of Otolaryngology at the University of Minnesota's Medical School, and co-director of the department's Lions Voice Clinic. There she treats patients with a wide range of voice disorders, specializing in care for professional singers. She also co-directs the voice research and education programs. She is a member and former chair of the Voice Science Advisory Committee for NATS, and the first moderator for the Vocapedia website. Dr. Michael continues to teach both voice and piano, and performs in a variety of venues.