In-Person or Virtual? − Assessing the Impact of COVID-19 on the Teaching Habits of Voice Pedagogues

*,†Kristen Murdaugh, *Josipa Bainac Hausknecht, and †Christian T. Herbst, *Salzburg, and †Vienna, Austria

Summary: The social distancing measures implemented world-wide in the wake of the novel Coronavirus (COVID-19) crisis have forced voice pedagogues to alter their teaching habits, likely shifting from customary in-person teaching to virtual teaching. An online survey, distributed world-wide in April/May 2020, investigated how singing voice pedagogues were impacted by the COVID-19 crisis. The collected responses from 387 survey participants suggest that, overall, voice teachers were only moderately satisfied with having to teach virtually, indicating that virtual voice teaching is not a sufficient replacement for in-person teaching. The participants indicated that during virtual teaching the singing voice can be assessed relatively well through features which provide both acoustic and visual clues. In contrast, depending on utilized technology, it may be harder to judge those aspects of the singing voice that are solely defined acoustically, such as dynamic range and spectral composition. This may be explained by limitations imposed by "out of the box" technology for online communication, which is typically optimized for speech instead of singing. This calls for better information on technological solutions for virtual voice teaching.
Key Words: Voice pedagogy−COVID-19−Online teaching−Voice transmission technology.

INTRODUCTION

Since its emergence in December of 2019, the novel Coronavirus (COVID-19) has heavily impacted both the global economy and society as a whole.[1] With early modeling studies indicating the need for intense control measures to limit the spread and severity of COVID-19 as well as to flatten the epidemic curve to reduce pressure on the healthcare system, social distancing measures were put in place around the world, having pronounced effects on professional life.[2]

Social distancing likely required many voice pedagogues of universities, as well as private and commercial studios, to change their teaching habits, shifting from customary in-person teaching to virtual teaching (In the context of this manuscript, customary in-person teaching refers to a teacher and student being in the same physical space, whereas virtual teaching refers to all teaching which does not take place in a shared physical space, but is rather facilitated by internet or telephone connectivity, using computers, mobile phones, or tablets as communication devices.). The shift from customary in-person to virtual teaching was particularly vital as societal lockdowns were implemented.

Addressing these fundamental changes of the professional landscape, this study aimed to document how voice teachers world-wide modified their teaching habits during the COVID-19-related lockdowns, as well as how content they were with the newly arising teaching situations and utilized technologies for virtual teaching. One particular section of questions targeted the individual attributes of the singing voice, and how well these can − in the opinion of the voice teachers − be assessed and evaluated through virtual teaching using video-conferencing tools like Zoom or Skype.

MATERIALS AND METHODS

Survey design

This study's data were acquired by means of an online survey. This "Virtual Voice Teaching Survey," available as a PDF in the supplementary materials, was constructed on the SurveyMonkey platform (SurveyMonkey Europe UC, Dublin, Ireland). It had the following structure: (1) Welcome (survey introduction and purpose; data protection declaration); (2) Demographic Data; (3) "Your Voice Teaching" (selecting the primary type, focus, style, and size of a pedagogue's studio); (4) Virtual Voice Teaching Technology (assessed with continuous sliders); (5) Assessment Criteria of the Singing Voice (assessed with continuous sliders); and (6) Closing (anonymous ID creation for future survey participation, optional final comment(s), statement of gratitude).

Every question without displayed answer options had a continuous sliding scale (visual analog scale, VAS) for assessment, with minima and maxima at 0 and 100, respectively. Instructions and notes were provided where relevant (see supplementary material). For example, the following note proceeded the Virtual Voice Teaching Technology page: "*the term 'virtual voice teaching/lesson' describes all measures that do not take place in the customary in-person setting (ie, student and teacher in the same room).*"

Voice assessment terms

One goal of this study was to investigate to what extent voice teachers can assess the individual attributes of the singing voice through virtual teaching using communication technology. This required a list of terms to assess the singing

voice. Surprisingly, despite scientifically informed suggestions for voice assessment terminology,[3] there does not yet seem to be an agreement as to which voice features are assessable in the singing voice via customary in-person teaching, let alone virtual teaching. Examining voice assessment terms listed in voice literature,[4−7] inconsistencies emerge, particularly with authors from differing backgrounds (ie, vocal pedagogy, voice therapy, acoustics) using differing vocabularies to describe voice quality.[3]

For these reasons, special care had to be taken to arrive at a comprehensive and representative list of voice assessment terms used in the "Assessment of the Singing Voice" section of the survey. Aiming to reflect the current standards of teaching institutions, this list was compiled as follows: the websites of North American English-speaking, postsecondary academic music institutions were reviewed. In particular, the institutions' websites were searched for voice evaluation rubrics, voice handbooks, and voice curricula containing voice assessment terms. Taking into consideration all singing styles, a cumulative table of 186 terms, found among 26 sources, was collected. In order to avoid the more obvious redundancies, all terms with a common initial word usage were semantically grouped, with the common word becoming the synonym for all others. An overview of the grouped terms is provided in the supplementary materials (note that only the terms which would otherwise result in redundancies due to their common initial word usage were grouped; nonredundant terms were considered "as is"). Grouping was performed algorithmically with a custom script written in Python by author CTH, reducing the list to 109 terms. The 15 most used terms were identified (see Figure 1) and listed in the survey. All grouped terms were listed in parenthesis next to the group synonym.

Possible redundancies were noted, and participants were asked to use their discretion to evaluate each term.

The survey's "Assessment Criteria of the Singing Voice" section requested that participants assess each term in the realm of virtual teaching with the following boundaries: *0 = This criterion cannot be assessed at all in the context of virtual singing teaching; 100 = This criterion can be fully assessed in the context of virtual singing teaching.*

Survey distribution and analysis

Ethical clearance for the survey was given by an *ad hoc* committee of the Mozarteum University's rectorate. The survey was distributed world-wide by email to voice organizations, music institutions, and colleagues of the authors. It was also posted in online voice teacher forums. Responses were collected for 1 month, beginning April 16, 2020.

The results were preliminary sighted with SurveyMonkey's analysis tools (question summaries, insights and data trends, individual responses). Data were subsequently analyzed using Python scripts written by CTH.

RESULTS

After one month of survey data collection, a total of 509 survey responses were recorded. 387 participants fully completed the survey, resulting in a 76.0% survey completion rate. The average response time was 6:35 [min:sec], with 5th and 95th percentiles at 3:37 and 27:08, respectively. Of all, 47.64% participants used a computer to complete the survey, 41.88% used a mobile phone, 10.21% used a tablet, and 0.26% used an alternate device. At the time of taking the survey, the participants had primarily taught virtually for an average of 34 days, with 5th and 95th percentiles at
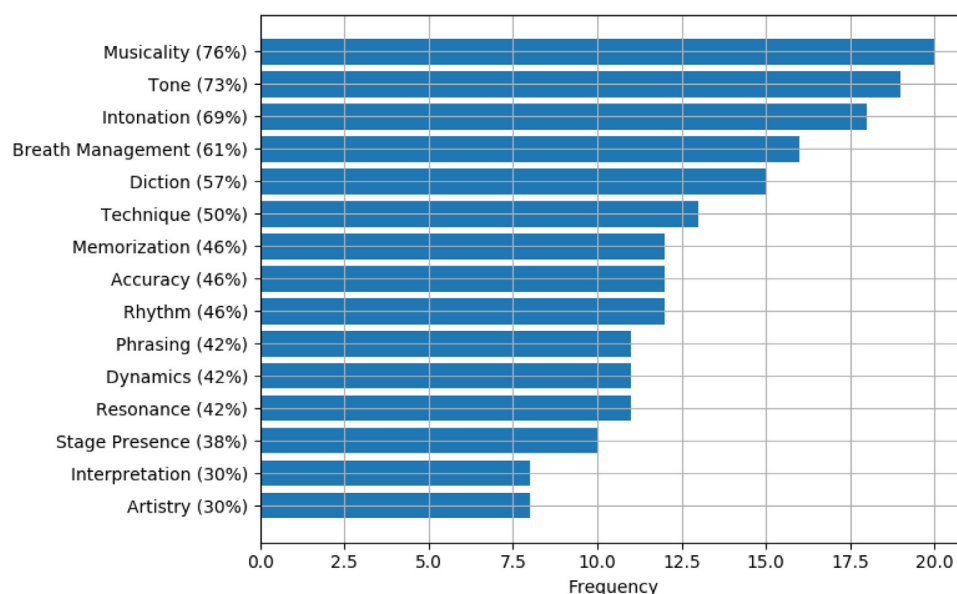
FIGURE 1. The 15 most commonly used voice assessment terms, including groupings, as found online in the voice evaluation rubrics, handbooks, and curricula of 26 North American English-speaking, postsecondary academic music institutions. The percentages provided in brackets indicate the frequency at which a particular term/term group was found in the 26 examined sources.
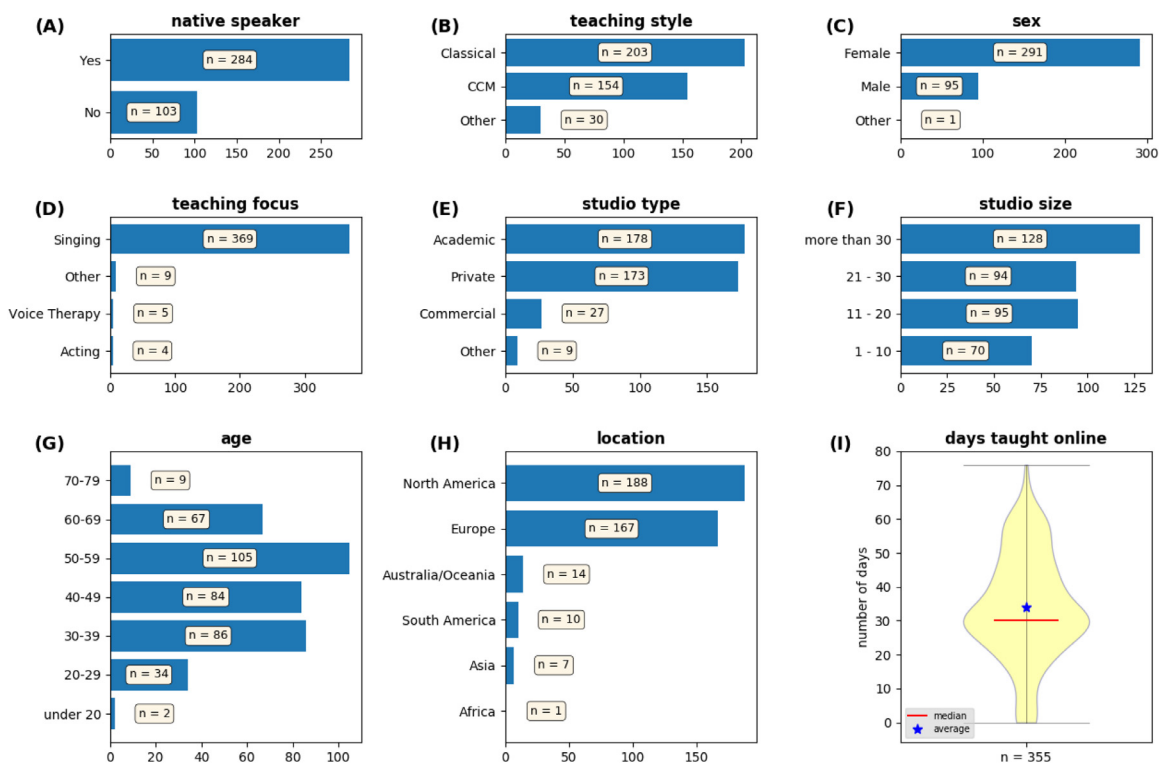
FIGURE 2. Demographic and voice studio data collected in the Virtual Voice Teaching Survey. The violin plot in panel I (days taught online) shows the average (mean) of the collected data with a blue star, the median of the collected data with a red line, and the distribution of the data through the varying width of the plot as well as the upper and lower horizontal lines which note the highest and lowest data points. Please note, the violin plot in panel I does not cover those survey participants (n = 29) who indicated that they "primarily taught online" prior to the COVID-19 crisis.

5 days and 60 days. Twenty-five participants (7.5 %) indicated that they primarily taught virtually prior to the COVID-19 crisis.

Analyzing the demographic data (see Figure 2), the prototypical survey participant was a 50−59 year old female native English speaker, who teaches more than 30 students per month in a classical singing-focused academic or private studio, located in North America or Europe, with Zoom being the primarily used virtual teaching technology.

As shown in Figure 3, voice teaching during the COVID-19 crisis is seen as important, with the majority of teaching having been converted to a virtual medium. As for the virtual medium used, online meetings with video and audio greatly surpassed the use of phone calls with only audio, while prerecorded sound clips were also somewhat of importance. Overall, teachers were averagely satisfied with teaching virtually and believed their students to be slightly more satisfied. However, there was a general consensus that
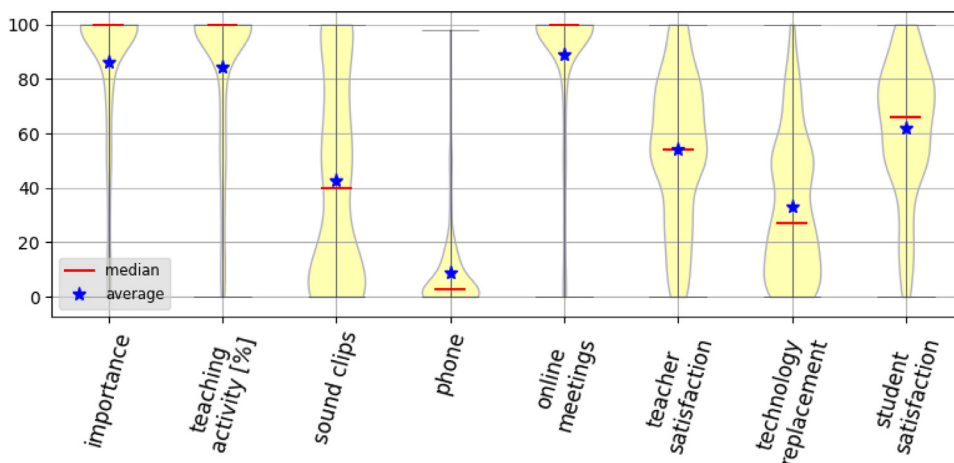
FIGURE 3. Survey participants' responses regarding virtual voice teaching and the related technologies.
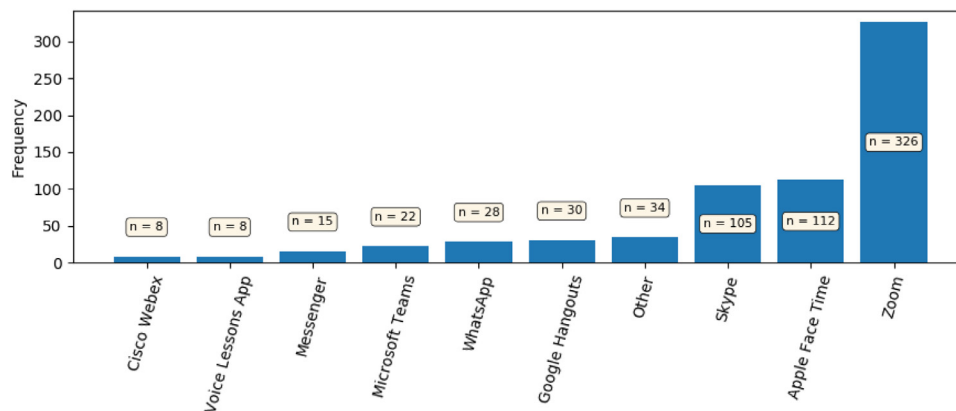
FIGURE 4. Primary technology used for virtual voice teaching by survey participants. Multiple replies were allowed for this question, resulting in a total of 668 responses provided by the 387 participants who completed the survey.

virtual voice lessons cannot replace customary in-person voice lessons, though the participants' opinions greatly differed in regards to this question (see the wide distribution of values in the "technology replacement" category of Figure 3).

On average, each survey participant (387 completed surveys) was using nearly two online technology platforms (668 responses collected − see Figure 4). Zoom was selected as the most prevalent virtual teaching technology, being used by almost half of the survey participants. Runners up were Apple's FaceTime and Skype, each being utilized by about a fourth of the survey participants.

In assessing the 15 provided voice assessment terms, participants' opinions greatly differed, as is shown in Figure 5. Keeping the assessment boundaries in mind (0 = *This criterion cannot be assessed at all in the context of virtual singing teaching; 100 = This criterion can be fully assessed in the context of virtual singing teaching*), every voice assessment term, with the exception of Rhythm and Interpretation, received assessments from 0 to 100, with averages and medians often differing greatly (hinting at non-normal distribution of the responses). With Memorization receiving an average rate of 81, Rhythm 76, and Diction 74, they were the 3 terms deemed most assessable in the context of virtual singing teaching. Most notably, the terms Dynamics, Resonance,

and Tone received the worst ratings, suggesting that overall the survey participants faced the greatest difficulties when assessing these voice features during virtual teaching. Interestingly, these three features are solely based on acoustic input and not reliant on visual information.

In order to better understand whether the possibility for assessment of these purely acoustic voice features (ie, Dynamics, Resonance, and Tone) is potentially linked to the use of certain virtual technologies, they were further examined in relation to the participants' primarily used technologies. The respective results are shown in Figure 6. Surprisingly, Messenger consistently had the best performance in regards to these three categories. However, Messenger was only used by 15 survey participants, with the typical user being a 40−49 year old female native English speaker who teaches 11−30 students per month in a classically or CCM singing-focused academic or private studio, located in North America. Notably, this participant profile significantly differs from that of the prototypical survey participant described above.

Similar to the survey data, final survey comments differed greatly. Many comments touched on the inability for technology to fully capture the singing experience, but also noted technology's ability to aid singers in developing their stage presence and online performance skills. Others
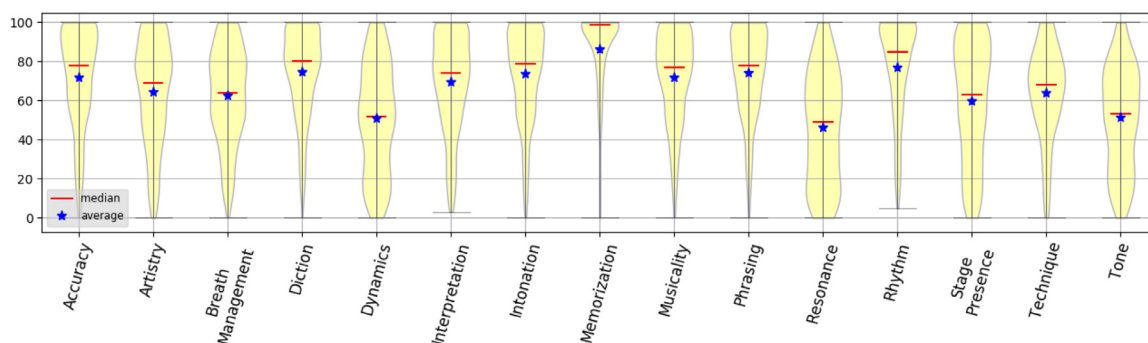
FIGURE 5. Survey participants' opinions on the virtual teaching assess-ability of the selected 15 voice assessment terms (n = 387 for all categories).
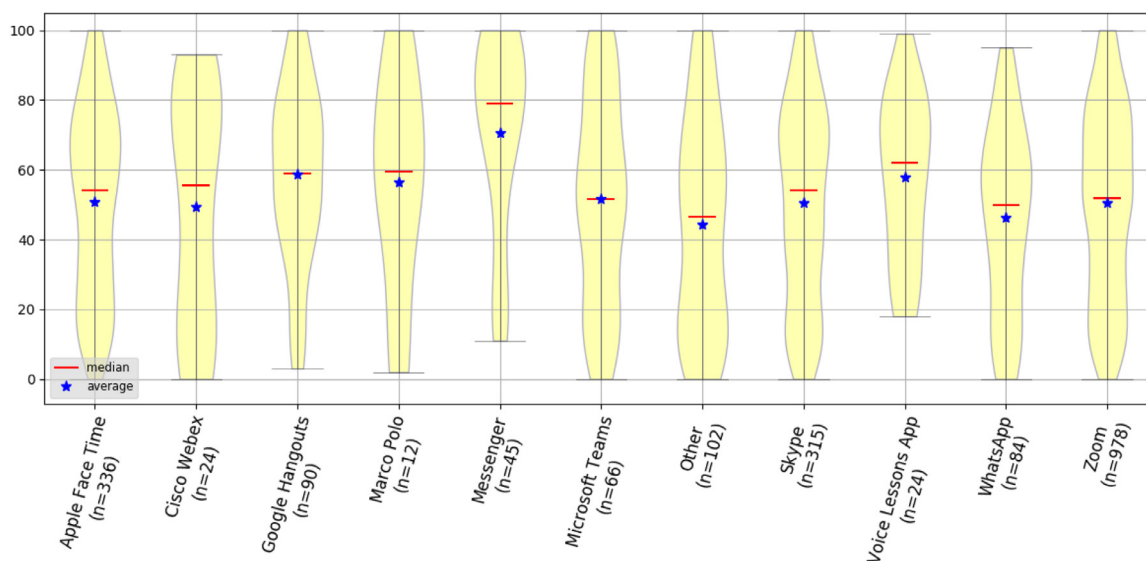
FIGURE 6. Survey participants' assessment of the purely acoustic aspects of the singing voice averaged over all three terms (ie, Dynamics, Resonance, Tone), and grouped by technology used.

remarked on the teacher and student need for better technology and more technological knowledge, particularly in regard to latency and data security. Several participants even commented on the joy of the challenge – learning to rely on more than visual cues for teaching, and instead honing the ear for acoustic variations. Overall, participants expressed gratitude that teaching may continue, but voiced discontent with the technology used to do so and the fatigue that using it may cause.

DISCUSSION

During COVID-19-related measures of social distancing, virtual voice teaching has become a central aspect of voice pedagogues' work. With the reality that the situation may continue for some time, it is important to understand what technology is currently being used and how satisfied teachers and students are with the available technology as well as the situation as a whole.

Overall, most voice teachers contributing to this survey seem to be only moderately satisfied with having to teach virtually, indicating that virtual teaching is not currently viewed as a sufficient replacement for in-person teaching. This notion may be comparable to online teaching and learning, ie, pedagogical concepts that have existed for more than 60 years,[8] with significant research and advancements in the last 20 years.[9] Recent studies suggest that students prefer, and perceive themselves to learn better through, online learning.[10] Yet, there still remains a divide between those who embrace and those who resist moving from in-person to virtual instruction. With the online teacher's role including pedagogical, social, managerial, and technical tasks, teaching online inevitably requires a specific set of skills which teachers may need to learn before they feel comfortable with virtual instruction.[11] However, with the rapid shift from in-person to virtual instruction due to

COVID-19, many voice teachers are left without this specific set of skills and potentially without the financial means to foster such skills.

The fact that, on average, each survey participant utilized almost two technological solutions (recall Figure 4) may suggest that – at least at this early stage of forced virtual teaching – voice pedagogues are still exploring options and technological solutions that would work for them. This hints at some potential parallels to the field of speech language pathology (SLP): In SLP, virtual telehealth services were also initially resisted due to various infrastructure shortcomings such as privacy and state licensure laws and technological deficiencies, but through a combination of synchronous and asynchronous technologies and years of technological research and developments in the SLP telehealth domain, clinical assessments and treatments can now be delivered virtually on a level similar to that of conventional in-person care.[12,13] Relying on computers, web-cameras, microphones, and internet access, with a focus on technologies which do not compromise the acoustic integrity of the transmitted voice signals, virtual SLP care became more equitable and accessible, particularly when asynchronous technologies were implemented as solutions to synchronous variabilities.[12] With this history in mind, SLP telehealth research may serve as a guiding source for voice teachers as they explore technological options and solutions in the COVID-19 crisis.

One central finding of this study was that voice teachers had greater difficulties rating most aspects of the voice that are solely grounded in acoustic cues, particularly dynamic range ("Dynamics") and spectral aspects ("Resonance," "Tone"). In contrast, other voice attributes which can be judged visually typically received better ratings. Surprisingly, the assertion differs as a function of technology used. While the majority of participants use Zoom, Apple Face-Time, and Skype, they only have an average satisfaction with

the acoustic features of these platforms, whereas users of Messenger express a higher satisfaction with the purely acoustic features. Formal and rigorous evaluation of these tools is required to identify whether qualitative differences exist among them. In support of this, a preliminary report provided by Howell et al[14] showed that each of four examined video conferencing platforms (Zoom, Microsoft Teams, VoiceLessonsApp, and FaceTime) negatively affect the transmitted voice signal at various but considerable degrees, highlighting the need for online communication systems that are better suited for the singing voice, considering issues like dynamic range, spectral composition and transmission lag.

That latter attribute (lag) points to another inherent issue, ie, internet broadband speed and latency. Caused by the pandemic-induced rise in virtual activity, networks have been severely impacted, with standard deviations of latency being nearly three to six times higher than those prior to the pandemic.[15] Interestingly, higher delays occur in the evening hours rather than the daytime working hours. This could have implications for the virtual voice teaching field, influencing the scheduling of lessons for optimal network quality.

Another aspect is that of sound input (microphones − see[16] for an excellent discussion) and sound output (loudspeakers/ earphones), both of which are key components in how acoustic features can be assessed in a virtual voice lesson. High-quality sound recording and playback requires high-quality equipment, which may be unaffordable to some. Furthermore, choosing and setting up the proper equipment (see[17] for a recommendable, if only slightly outdated tutorial) may pose technical challenges. While it may seem obvious that − for purposes of virtual voice teaching − the entire technological voice transmission chain (sound recording, transmission, and playback) needs to truthfully and realistically represent the entire frequency and dynamic range of the singing rather than the speaking voice, this criterion is not always met. Rather, some components are designed with the improved transmission of the speaking voice in mind, implicitly and automatically "improving" the signal (which, in fact, may introduce unwanted distortions to the transmitted singing voice signal). For instance, some microphones may under certain circumstances add a "bass boost" or a boost to the singers' formant region between 2 and 4 kHz,[16] and some loudspeakers and earphones may alter the spectral or dynamic composition of the sound via a feature called active noise control (ANC) in an attempt to equalize and attenuate ambient sound.[18]

This study has been designed as an *ad hoc* response to the emerging COVID-19 crisis. As such, it may suffer from a number of potential limitations:

For practical reasons, only voice assessment resources from North American English-speaking universities have been considered. Research shows that using a common language (such as English) rather than native languages in cross-national questionnaires can obscure cultural differences in the resulting data.[19] However, using a common language for cross-national research is also recommended to avoid translation errors and resulting data discrepancies.[20]

Another issue might come from a potential limitation to the list of voice assessment terms (recall Figure 1), since the data pool was narrowed down to available online resources from 26 North American universities. Given the surprising number of different voice assessment terms found in these 26 sources (186 individual terms, reduced to 109 terms through semantic grouping − see Methods), a larger formal study is warranted, collecting and reviewing voice assessment terms utilized in English-speaking tertiary education institutions world-wide.

Finally, the survey results may have been influenced by the choice to use continuous (VAS-like) sliders over discrete scales (Likert scale) for term assessment. Studies comparing continuous (VAS-like) sliders with discrete evaluation scales, such as the Likert scale, show positives and negatives for both options.[21] While discrete evaluation scales may evoke data inconsistencies due to the effect of response-order, continuous sliders may suffer from functionality issues leading to low response rates. That latter aspect may explain this survey's relatively low completion rate of 76%.

SUMMARY

The data from this survey suggest that singing voice teachers world-wide have been forced to convert their teaching activity to virtual teaching with little time for preparation of particularly the technological aspects of that transition. Therefore, there may be a great potential for improvements of current teaching situations. There is a need for (1) clearer information on the technical capabilities of systems/setups; (2) recommendation(s) backed up by a consensus group of "tech-savvy voice teachers"; and (3) the means to provide this information and the required technical know-how to the international community of voice teachers.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at doi:10.1016/j.jvoice.2020.08.027.

REFERENCES

1. Sohrabi, C, Z Alsafi, N O'Neill, et al. 2020. World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). 76: 71−76. https://doi.org/10.1016/j.ijsu.2020.02.034.

2. Prem K, Liu Y, Russell TW, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *Lancet Public Health*. 2020;5:e261–e270. https://doi.org/10.1016/S2468-2667(20)30073-6.

3. H Bernardoni, Nathalie, P Bezard, et al. 2008. Towards a common terminology to describe voice quality in western lyrical singing: contribution of a multidisciplinary research group. 2. Vol. 2. Gordon; Breach. https://hal.archives-ouvertes.fr/hal-00297248.

4. Vennard W. *Singing: The Mechanism and the Technic.*. Carl Fischer; 1968.

5. Miller R. *The Structure of Singing: System and Art in Vocal Technique*. Schirmer Books; 1986.

6. Guerin, C. 2006. Glossaire de l'Atelier du Chanteur. https://chanteur.net/glossaire.htm.

7. Garnier M, Yargₗlarₗndan E, Betimleyicilere A, et al. Characterisation of voice quality in western lyrical singing: from teachers' judgements to acoustic descriptions. *J Interdiscip Music Stud*. 2007;1:62–91.

8. Nicholson P. A history of E-learning: echoes of the pioneers. *Computers and Education: E-Learning, from Theory to Practice*. Netherlands: Springer; 2008:1–11. https://doi.org/10.1007/978-1-4020-4914-9_1.

9. Palloff RM, Pratt K. *Lessons from the Virtual Classroom: The Realities of Online Teaching*. 2nd Edition Wiley. Second. Jossey-Bass; 2013. https://www.wiley.com/en-us/Lessons+from+the+Virtual+Classroom{\%}3A+The+Realities+of+Online+Teaching{\%}2C+2nd+Edition-p-9781118123737.

10. Hannay M, Newvine T. Perceptions of distance learning: a comparison of online and traditional learning. *Merlot J Online Learn Teach*. 2006;2. http://www.assoa.nt.edu.au/how{_}it{_}began.html.

11. Keengwe J, Kidd TT. Towards best practice in online learning and teaching in higher education. *J Online Learn Teach*. 2010;6:533–541.

12. Keck CS, Doarn CR. Telehealth technology applications in speech-language pathology. *Telemed E-Health*. 2014;20:653–659. https://doi.org/10.1089/tmj.2013.0295.

13. Mashima, PA, DP Birkmire-Peters, MJ Syms, et al. 2003. Telehealth: voice therapy using telecommunications technology. https://doi.org/10.1044/1058-0360(2003/089).

14. Howell, IL, KJ Gautereaux, J Glasner, et al. 2020. Preliminary report: comparing the audio quality of classical music lessons over zoom, Microsoft teams, VoiceLessonsApp, and Apple FaceTime. https://www.ianhowellcountertenor.com/preliminary-report-testing-video-conferencing-platforms.

15. Candela, M, V Luconi, A Vecchio. 2020. Impact of the COVID-19 pandemic on the internet latency: a large-scale study, May. http://arxiv.org/abs/2005.06127.

16. JG S, Granqvist S. Guidelines for selecting microphones for human voice production research. *Am J Speech-Lang Pathol*. 2010;19. https://doi.org/10.1044/1058-0360(2010/09-0091).

17. Spielman J, Starr AC, Popolo PS, et al. *Recommendations for the Creation of a Voice Acoustics Laboratory*. Denver, CO: National Center for Voice; Speech; 2007. http://www.ncvs.org/ncvs/library/tech/NCVSOnlineTechnicalMemo07.pdf.

18. Fabry J, Konig F, Liebich S, et al. Acoustic equalization for headphones using a fixed feed-forward filter. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, *2019-May:980−84*. Institute of Electrical; Electronics Engineers Inc; 2019. https://doi.org/10.1109/ICASSP.2019.8682167.

19. Harzing A-W. Does the use of english-language questionnaires in cross-national research obscure national differences? *Int J Cross Cult Manag*. 2005;5:213–224. https://doi.org/10.1177/1470595805054494.

20. Sechrest L, Fay TL, Hafeez Zaidi SM. Problems of translation in cross-cultural research. *J Cross-Cult Psychol*. 1972;3:41–56. https://doi.org/10.1177/002202217200300103.

21. Chyung SYY, Swanson I, Hankinson Andrea, et al. Evidence-based survey design: the use of continuous rating scales in surveys. *Perform Improv*. 2018;57:38–48. https://doi.org/10.1002/pfi.21763.