Show and Tell: Video Modeling and Instruction Without Feedback Improves Performance but Is Not Sufficient for Retention of a Complex Voice Motor Skill

Clarisse Look, Patricia McCabe, Robert Heard, and Catherine J. Madill, *Lidcombe, New South Wales, Australia*

Summary: Purpose. Modeling and instruction are frequent components of both traditional and technology-assisted voice therapy. This study investigated the value of video modeling and instruction in the early acquisition and short-term retention of a complex voice task without external feedback.

Method. Thirty participants were randomized to two conditions and trained to produce a vocal siren over 40 trials. One group received a model and verbal instructions, the other group received a model only. Sirens were analyzed for phonation time, vocal intensity, cepstral peak prominence, peak-to-peak time, and root-mean-square error at five time points.

Results. The model and instruction group showed significant improvement on more outcome measures than the model-only group. There was an interaction effect for vocal intensity, which showed that instructions facilitated greater improvement when they were first introduced. However, neither group reproduced the model's siren performance across all parameters or retained the skill 1 day later.

Conclusions. Providing verbal instruction with a model appears more beneficial than providing a model only in the prepractice phase of acquiring a complex voice skill. Improved performance was observed; however, the higher level of performance was not retained after 40 trials in both conditions. Other prepractice variables may need to be considered. Findings have implications for traditional and technology-assisted voice therapy.

Key Words: Voice therapy–Modeling–Instruction–Motor learning–Acquisition–Prepractice.

Nearly 30% of adults experience a voice disorder in their lifetime.[1] Voice therapy traditionally involves regular sessions with a speech-language pathologist and a prescribed home program. Direct voice therapy aims to establish new motor patterns or re-establish previously healthy vocal movement routines.[2] To do this, voice clinicians use a number of elements including modeling, instruction, and feedback to assist the acquisition and retention of vocal skill.[2] This process involves motor learning. Consequently, in recent years, there has been increasing scientific interest in the application of principles of motor learning (PML) to voice therapy.[3–5] Despite this, the use of PML to enhance vocal task learning, particularly in the early or prepractice phase of therapy, is still not well understood. Therefore, this paper investigates the effect of two therapy behaviors—modeling and instruction—on the acquisition and short-term retention of a complex voice task.

PRINCIPLES OF MOTOR LEARNING

Motor learning refers to the acquisition, retention, and transfer of skilled movements over time.[6] Acquisition has been described as occurring over two phases: prepractice and practice.[7] Motor learning research uses the term prepractice to describe information provided to learners before task attempts.[8,9] According to these researchers, practice begins once learners attempt the task. However, speech-language pathology (SLP) often regards prepractice as a phase that includes initial task attempts. This paper will use Maas et al's[7] definition of prepractice, which states that during this phase, clients are guided to understand the task, establish an internal reference of correctness, and perform the task under optimal conditions. When the prepractice goals are achieved, clients move on to the practice phase. In this phase the target skill is repeated, using selected PML, to complete acquisition and to facilitate both transfer and maintenance of performance. In motor speech studies, 80% accuracy over two to three sessions or blocks is commonly used as the acquisition criterion.[10–12] It is important to note that performance during acquisition does not predict learning, as learning is only achieved when there is both skill retention and transfer.[7]

THE IMPORTANCE OF PREPRACTICE

As illustrated by two systematic reviews,[7,13] most PML research in SLP has centered on the practice phase. In comparison, prepractice is often overlooked. There is some evidence for the effectiveness of discovery learning, which is when little prepractice information is provided and learners trial various strategies to acquire a task.[14] However, discovery learning is not recommended for voice, as incorrect hyperfunctional practice may lead to further disorder.[15] The prepractice phase is therefore essential in ensuring that practice is both healthy and beneficial. In addition, prepractice is typically conducted with a clinician during face-to-face clinic or telehealth sessions, and represents the primary cost in SLP intervention. Most practice is completed outside the clinic setting, usually in the form of home exercises. Maximizing the efficiency of prepractice through research will thus maximize therapy time and resources, leading to improvements in cost-effective service provision. In voice therapy, modeling, instruction, feedback, motivation, and perceptual training are key prepractice variables.[2]

Accepted for publication September 27, 2017.

From the University of Sydney, Lidcombe, New South Wales, Australia.

Address correspondence and reprint requests to Catherine J. Madill, Discipline of Speech Pathology, Faculty of Health Sciences, University of Sydney, PO Box 170, Lidcombe, NSW 1825, Australia. E-mail: cate.madill@sydney.edu.au

Journal of Voice, Vol. 33, No. 2, pp. 239–249
0892-1997

https://doi.org/10.1016/j.jvoice.2017.09.023

GENERAL MOTOR PLANS

To achieve optimal motor learning, use of the PML in therapy facilitates the generation and consolidation of a generalized motor program (GMP) over repeated trials. According to Schema Theory, this GMP defines invariant features of the target movement. This GMP is then modified by varying a range of parameters, which specify how the program is executed in different conditions. Hence, when different parameters are applied, the same GMP can produce multiple variations of the target movement.[6,16] For example, whereas the GMP represents the relative sequence, timing, duration, order, and force of muscle contractions, parameters determine the absolute values of these features according to task demands. The process of establishing and refining GMPs occurs throughout prepractice and practice in motor learning.[6]

MODELING AND INSTRUCTION

Modeling uses an external stimulus to provide information about subtle skill components that can be difficult to teach with verbal instruction. When clients observe a model, they gain a cognitive impression of the target form, which promotes effective learning of that skill.[17,18] The use of models is well supported in the motor learning literature (see Hodges and colleagues[19] for review). Research shows that a combination of observation and physical practice leads to enhanced learning, and that learners engage in better practice after several observations during prepractice and early practice.[20,21] There is also evidence of neurological activation in the frontal and parietal brain regions when someone watches a model.[22] Such activation contributes to enhanced motor representations, as well as improved perceptual abilities in predicting movement outcomes.[23]

Instruction on the other hand can either facilitate or inhibit learning. It contributes to the development of a reference of correctness, resulting in greater understanding of the target skill.[24] Instruction also plays a role in directing focus of attention—attending to movement effects, rather than movement process, is more beneficial to learning.[25] This is related to promoting a more implicit form of learning that is relatively robust to factors such as pressure, fatigue, and time.[26,27] However, detailed instructions can impede learning by overloading the cognitive system and preventing active engagement in the learning process.[28,29] This may be especially true for complex skills.[29] It is important to explore the use of instruction in voice therapy, as research has shown that voice clinicians engage in significantly more talking behaviors than doing behaviors during sessions.[30]

Task complexity is one of the many factors that influence the effectiveness of modeling and instruction.[8] Complexity relates to the cognitive load required for a task, which is associated with the number of task elements.[31] Wulf and Shea[32] proposed that facilitatory conditions for simple skill learning might not apply to complex skills. They postulated that simple skills involving low cognitive load benefit from more challenging conditions. Conversely, the learning of complex skills is enhanced when processing demands are reduced to manageable levels. This notion is also supported by the *challenge point framework*, which states that learning is maximized when learners are optimally challenged.[33] Thus, task complexity is an important consideration in motor learning studies.

GENERAL MOTOR TASKS VERSUS VOICE TASKS

Notably, available research relating to modeling and instruction is based on general motor skills, which involve whole body or limb movements. A recent speech review suggested that principles derived from general motor tasks might not directly relate to other systems.[13] There are several reasons why this might be true for voice. First, when modeling a general motor task, learners can visualize the process of achieving the target outcome.[19] For example, they can observe arm movement for a tennis serve. In contrast, voice models only highlight the outcome as clients cannot view the workings of the larynx. The lack of this visual information during voice task modeling may affect the effectiveness of this prepractice variable. Second, there are differences in task characteristics. Many motor learning studies are based on high-performance tasks for specific contexts such as sport,[19] whereas spoken voice is an unconscious, everyday activity. Finally, many motor learning studies investigate the learning of a new task.[18,21] In contrast, voice therapy often aims to replace a disordered motor pattern with a previously established appropriate pattern.[2]

VOICE THERAPY AND TECHNOLOGY

In other areas of health care, technology is widely used to improve service efficiency and quality[34] and increasingly, clinicians and researchers have started to explore the use of technology-assisted therapy in voice.[35] Technology has brought about a proliferation of voice teaching material in the form of computer programs and mobile applications, or apps. These are available in both audio only and audiovisual forms. For example, there are many voice apps on the Apple Store and Google Play (eg, VoixTek Voice Training,[36] Vocal Ease[37]) and an even greater number of videos online (eg, The Hum,[38] Vocal Straw Exercise[39]). With the dramatic rise in technology use, voice clinicians can expect mobile and computer tools to become increasingly popular.

Many voice computer and mobile programs are designed for home-based practice or to supplement traditional face-to-face therapy.[35] Several prepractice variables can be clearly identified in existing mobile and computer voice therapy and training tools. Many programs present a model demonstrating the target voice behavior and instructions providing information for skill acquisition. In addition, interactive programs measure task performance and provide online external feedback. A myriad of programs are specifically designed to provide such external feedback both in therapeutic and skill-acquisition settings about voice pitch, intensity, and quality, including the Lee Silverman Voice Treatment Companion,[40] Opera Slinger,[41] Sonneta Voice Monitor,[42] Sing and See,[43] Voice Analyst,[44] and OperaVOX.[45]

On the other hand, some programs do not include an interactive function; they are limited to the provision of a model with or without instruction in facilitating vocal skill acquisition. All online training videos such as those on YouTube and commercial instructional videos (eg, Voice Care for Teachers[46]) are noninteractive, as users do not receive external feedback. Many mobile apps follow a similar format. For example, Voice Builder[47] and Virtual Voice Lessons[48] are noninteractive apps that provide verbal instructions and video demonstrations for voice exercises. Although technology promotes more practice, it is not known if it has a positive effect on acquisition or retention of the targeted voice skill.

As voice training technologies become more accessible, it is increasingly important for clinicians to understand how these tools can assist acquisition and retention, and whether they can operate as independent, stand-alone treatments. As modeling and instruction are frequent prepractice variables in voice programs, closer examination of these components is warranted. To date, no study has investigated modeling and instruction in the context of voice therapy and training.

RESEARCH QUESTIONS

Despite considerable research in general motor learning, the role of PML in the voice therapy field is still not well understood. Modeling and instruction are two of the most commonly used prepractice variables; effective use of these variables will become increasingly significant as computer and mobile technology play a greater role in service delivery. Therefore, this study focused on the role of modeling and instruction in voice motor learning. Our specific research questions and hypotheses were as follows:

Research question 1: Is video modeling without external feedback sufficient to enable improved performance of a complex vocal task?

Hypothesis 1: The provision of a model with or without instruction is sufficient for the acquisition and short-term retention of a complex voice task.

Research question 2: Does verbal instruction significantly impede or assist performance of a complex motor task in the context of video modeling without external feedback?

Hypothesis 2: Providing an instruction with the model will be more beneficial than providing a model only.

METHOD

Participants

Thirty female participants aged 18–33 years ($M = 21.5$, $SD = 4.2$) were recruited from the SLP student population at The University of Sydney via posters, lecture announcements, and e-mail. This sample size was consistent with a similar voice study by Joscelyne-May and colleagues.[49] All participants spoke fluent English, had no history of voice disorder, and were unfamiliar with the target vocal task. Before the experimental protocol, participants also passed hearing[50] and voice[51] screenings. Exclusion criteria were previous vocal training, completion of any SLP voice course, and risk factors for voice disorder including advanced age, smoking, and regular inhaled corticosteroid use.[1,52,53] This study was approved by The University of Sydney Human Research Ethics Committee (2014/042). Each participant gave written consent before commencing the study.

Voice task

Participants were trained to produce a vocal siren on a nasal sound using optimal vocal production, that is, a clear and effortless voice. This was described to the subjects as "start with a nasal *ng* sound, and then gradually decrease and increase your pitch while maintaining a constant volume. Voice production should be clear and effortless."

A vocal siren is a continuous pitch glide up and down, most often produced on a nasal ("ng") sound. The vocal siren is designed to improve vocal range; however, Harris[54] points out that sirening "can help patients learn to gain control of the cricothyroid visor, allowing smooth changes in vibrating vocal fold mass in order to shift from thin to thin folds. It also encourages flexibility of vertical laryngeal position" (p. 180). Therefore, the siren is used generically as an exercise to increase control of multiple vocal features within a single task. Pitch glides are used extensively across common voice therapies such as Stemple's Vocal Function Exercises[55] and Lee Silverman Voice Technique.[56] The added task of maintaining volume constant was given to create greater complexity as the participants had vocal function within normal limits. A vocal siren was chosen as it is typically unfamiliar to individuals without formal voice training. Hence, this task should have been novel to all participants. Furthermore, a siren can be considered a complex task. Producing a vocal siren requires simultaneous regulation of many parameters and the task can be broken down into smaller components (ie, control of pitch, volume, voice quality, and respiration). Given that participants had healthy voices, a complex task was required to measure acquisition and retention. This task has been used in previous motor learning research with nondisordered subjects as a valid, measurable task sensitive to acquisition effects.[49]

Procedure

Participants attended two sessions: a training session and a retention session. These were conducted in the Voice Laboratory at the Communication Disorders Treatment and Research Clinic of the University of Sydney. Drinking water was available during all sessions.

All participants were randomly allocated to one of two treatment conditions: (a) model only or (b) model and instruction. The training protocol was scripted, video-recorded, and edited to produce separate training videos for each group. These videos showed the second author providing models with or without instructions, depending on the participants' allocated group. An expert model was chosen for ecological validity, as voice clinicians often model the target they want their client to acquire and learn.[30] The same model and instructions were provided throughout the video. After viewing of the relevant model (with or without instructions), participants were directed to attempt the task twice.

They viewed the videos at the same volume, on a 13-inch laptop monitor placed approximately 1 m away. During sessions, participants were fitted with head-mounted microphones placed at a measured 5 cm from their lips. Participants completed four blocks of 10 trials for a total of 40 attempts. This number of trials was chosen based on previous voice motor learning research by

TABLE 1.
Training Schedule

Group	Session 1 (Training)			Session 2 (Retention)
	Block 1	Block 2	Block 3	Block 4
Model-only group	Model only	Model only	Model only	Model only
Model and instruction group	Model only	Model + instruction	Model + instruction	Model only

Steinhauer and Grayhack.[4] Participants performed three blocks during the first session (training), and then performed one block during the second session (retention) held 24–30 hours later. Table 1 shows the training schedule for both groups over two sessions. The model-only group received a model for all blocks, whereas the model and instruction group received additional verbal instructions in the second and third block. Both groups saw the model once at the start of each block of training to hold the dose of exposure to the model consistent across participants and groups. The training instructions for the model and instruction group were "Start with an *ng* sound at your normal speaking pitch. Gradually increase and decrease your pitch. Repeat this slowly while maintaining a constant, quiet volume. Your voice should be clear and easy to produce." Both groups received a model only in the first block to allow for measurement of baseline equivalence, ensuring that there was no significant difference in learning speed between the groups under the same conditions. There was a 1.5-minute break between blocks. To prevent uncontrolled practice, participants were directed to engage in a word search puzzle or to drink water during the break.

Equipment

The sessions were audio-recorded in a soundproof booth using a C420^{III} cardioid microphone. Recordings were made with a Layla 24/96 Multitrack Recording System (Echo, Santa Barbara, CA) and Adobe Audition 1.5 (Adobe Systems Inc., San Jose, CA) at a sampling rate of 44 100 Hz and 16-bit resolution. They were saved in .wav format for editing with *Wavepad* (NCH Software, Greenwood Village, CO).[57]

Measures

The initial and final siren of the entire experiment, and the last sirens of each training block (ie, the 10th, 20th, 30th siren) were extracted using *Wavepad*.[57] These were labeled T1, T2, T3, T4, and T5, respectively. The first author selected the sirens by hand and extracted them using the Wavepad trimming function, automatically trimming silence from the start and end of the file. The extracted sirens were then analyzed for phonation time, vocal intensity, cepstral peak prominence (CPP), peak-to-peak time, and root-mean-square error (RMSE). These measures are key features of the siren, and hence improvements can show the process of early acquisition.

Phonation time and vocal intensity

For each siren, *Praat* software (Phonetics Sciences, Amsterdam)[58] was used to establish the phonation time in seconds and mean-energy vocal intensity in decibels. These measures are appropriate as phonation time reflects phonatory and respiratory function in voice,[59] whereas reduced vocal intensity is a feature of the vocal siren. As no instruction was given to subjects to match the length of the siren, this was considered a control measure to evaluate the effect of the number of trials, not the instruction condition. Vocal intensity was an instructed feature, so a reduction in vocal intensity would indicate early acquisition of that feature of the task.

Cepstral peak prominence

CPP correlates with the periodicity of the voice signal and can be used to distinguish between normal and dysphonic voices.[60] This measure is suitable for quantifying voice quality in vocal sirens, as calculation of CPP is not dependent on frequency tracking. For CPP analysis, a 4-second sample of every siren was obtained. To exclude voicing onset, siren samples were set for between 1 and 5 seconds in *Praat*.[58] CPP was then calculated using a windows command prompt.[61] It is known that CPP values will reduce with reduced vocal intensity. A reduction in CPP coinciding with a reduction in volume would be expected; however, maintenance of CPP values or an increase in CPP following a reduction in volume, would indicate an improved vocal quality, that is, more clarity in the voice, as instructed in the task.

Peak-to-peak time

Peak-to-peak time provided a measure of vocal control, reflecting the participants' ability to match the model while manipulating pitch in a regular, periodic manner. The Sing and See program[43] provided an accurate visual representation of each siren's pitch contour. Peak-to-peak time was defined as the distance between the first and second pitch peak. This distance was measured by hand, and the timing between peaks was automatically calculated in seconds. Manual detection of the peaks was necessary to exclude false peaks due to pitch breaks.

Root-mean-square error

RMSE is a measure of fit—it is typically used to reflect differences between predicted and actual values.[62] In this study, it was used to compare the pitch and timing characteristics of the siren attempts with the expert model. Whereas the other outcome measures mostly reflect characteristics at a specific section of the siren, RMSE accounts for the full siren recording. Hence, it can be considered a whole-task measure.

To determine RMSE, *Praat*[58] was used to generate pitch listings of the expert model and each siren attempt at 0.01-second intervals. The first and last 0.04 seconds of all samples were excluded from analyses. Next, the difference between the pitch of the expert model and the pitch of the participant's attempt for each time point was calculated to generate difference scores. A constant was subtracted from the difference scores to correct for variation in starting pitch. The standard deviation of the corrected difference scores was then calculated in Microsoft Excel 2011, providing the RMSE.

Reliability

An independent researcher replicated hand sampling and generation of scores for 10% of the data. Inter-rater reliability was calculated using two-way mixed model intraclass correlation coefficients ICC (3,1). An ICC value above 0.80 was considered excellent.[63] There were excellent ICC values for reliability of sample extraction for phonation time (1.00), intensity (1.00), CPP (1.00), peak-to-peak time (0.997), and RMSE (1.00).

Normality

The distribution of data was inspected for skewness and kurtosis using z-scores. All phonation time and intensity data met the assumption of normality. Thus, parametric tests were used for these measures. However, some data points for CPP, peak-to-peak time, and RMSE did not fit a normal distribution. Nonparametric tests were used for these analyses.

Baseline equivalence

Baseline equivalence was established by statistically comparing vocal siren performance between both groups at T1 and T2. This was achieved with a 2×2 analysis of variance (ANOVA) (parametric) or a Mann-Whitney U test comparing the median of T1 and T2 (nonparametric). Whereas nonparametric tests were used for CPP in the subsequent measures, parametric tests were used for CPP at baseline as the data at T1 and T2 were normally distributed.

At baseline, there were no significant differences ($P > 0.05$) between the performance of the groups on measures of phonation time, ($F(1, 28) = 3.309$, $P = 0.080$, partial $\eta^2 = 0.106$), vocal intensity, ($F(1, 28) = 0.177$, $P = 0.677$, partial $\eta^2 = 0.006$), CPP, ($F(1, 28) = 0.036$, $P = 0.851$, partial $\eta^2 = 0.001$), peak-to-peak time, ($U = 77$, $z = -1.473$, $P = 0.141$, $r = -0.269$), and RMSE ($U = 103$, $z = -0.394$, $P = 0.694$, $r = -0.0719$).

RESULTS

Performance of vocal siren over trials

To address our first research question, the model-only group, and the model and instruction group were analyzed separately to determine change in siren performance over trials. For phonation time and intensity, 1×5 repeated measures ANOVAs with standard and polynomial contrasts were used. For CPP, peak-to-peak time, and RMSE, Friedman tests were conducted with post hoc analyses using Wilcoxon signed-rank tests. We note that medians are typically reported for nonparametric statistics. However, as both parametric and nonparametric tests were used in this study, we used means in the figures for consistency and to facilitate interpretation. Figures 1 to 5 present the means at all time points for phonation time, vocal intensity, CPP, peak-to-peak time, and RMSE, respectively. The dashed lines in these figures represent values from analysis of the expert model's siren.

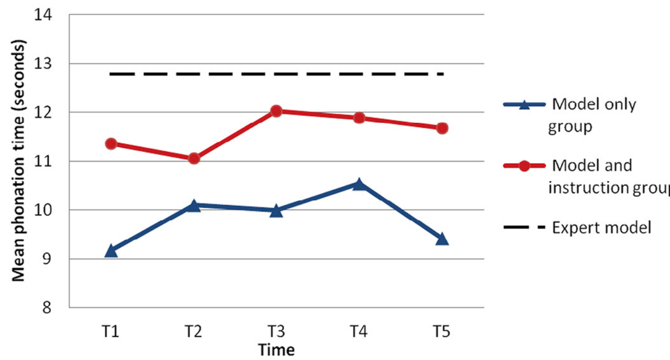

FIGURE 1. Mean phonation time over trials for model-only group, model and instruction group, and expert model.

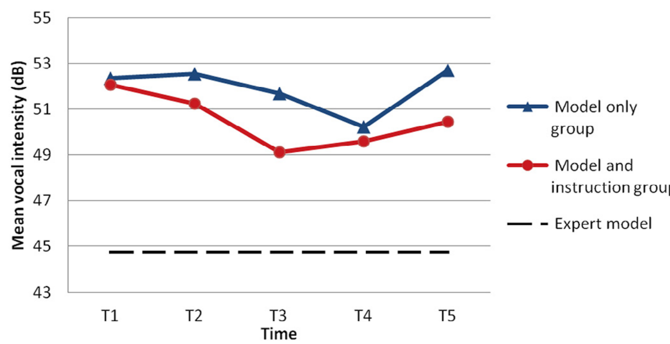

FIGURE 2. Mean vocal intensity over trials for model-only group, model and instruction group, and expert model.

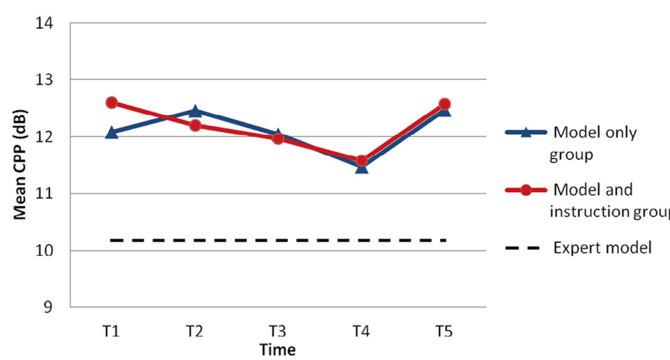

FIGURE 3. Mean CPP over trials for model-only group, model and instruction group, and expert model.

Significant quadratic change in performance over trials was observed in both the model-only group ($F(1, 11) = 5.797$, $P = 0.030$, partial $\eta^2 = 0.293$) and the model and instruction group ($F(4, 11) = 4.647$, $P = 0.019$, partial $\eta^2 = 0.628$). In the model-only group, an increase in mean phonation time toward the model from T1 to T4 (within session one), followed by a sharp decrease from T4 to T5 (from session one to session two), was observed. In the model and instruction group, mean phonation time increased from T1 to T3, and then gradually decreased from T3 to T5. Follow-up analysis showed that performance improved significantly from T2 to T3, $F(1, 11) = 7.331$, $P = 0.017$, partial $\eta^2 = 0.344$, when the instruction was first introduced.

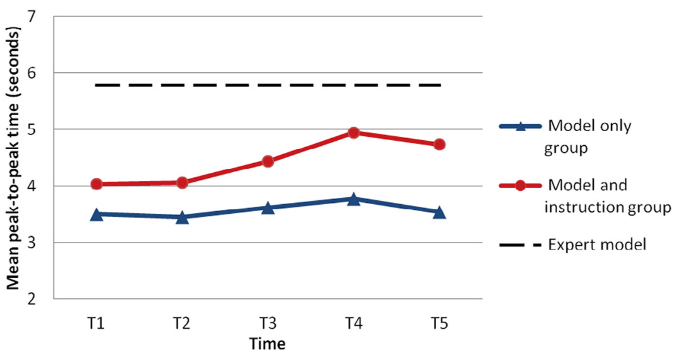

FIGURE 4. Mean peak-to-peak time over trials for model-only group, model and instruction group, and expert model.

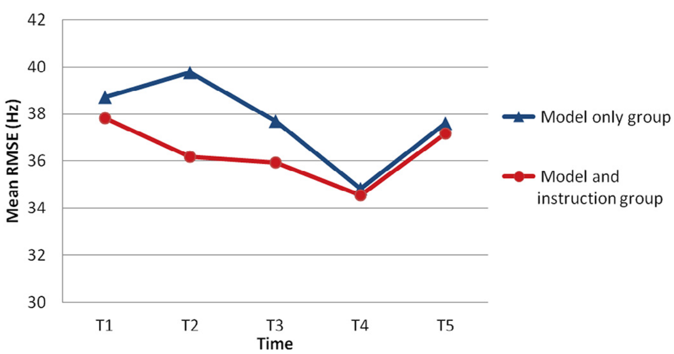

FIGURE 5. Mean RMSE over trials for model-only group, and model and instruction group.

Intensity measures revealed significant change over trials in both the model only group ($F(4, 11) = 9.138$, $P = .002$, partial $\eta^2 = .769$), and model and instruction group, ($F(4, 11) = 6.850$, $P = .005$, partial $\eta^2 = .714$). Further comparisons showed the model only group decreased towards the model in mean intensity from T1 to T4 and then increased from T4 to T5 (from session 1 to session 2), patterns revealed by a significant quadratic, ($F(1, 11) = 5.417$, $P = .035$, partial $\eta^2 = .279$) and cubic effect ($F(1, 11) = 37.831$, $P < .001$, partial $\eta^2 = .730$.) In the model and instruction group, mean intensity decreased in the first session from T1 to T3, and then increased gradually from T3 to T5, patterns revealed by a significant quadratic, ($F(1, 11) = 9.808$, $P = .007$, partial $\eta^2 = .412$), and order 4 effect, ($F(1, 11) = 4.677$, $P = .048$, partial $\eta^2 = .250$).

CPP measures in both the model-only group (Friedman $\chi^2(4) = 12.107$, $P = 0.017$) and the model and instruction group (Friedman $\chi^2(4) = 9.867$, $P = 0.043$) changed significantly over trials. In the model-only group, post hoc analysis found that CPP values decreased significantly toward the model between T2 and T4 (within session one), $z = -2.528$, $P = 0.011$, $r = 0.653$, and then increased significantly from T4 to T5 (from session one to session two), $z = -2.669$, $P = 0.008$, $r = 0.689$. In the model and instruction group, CPP values followed a similar trend: a significant decrease toward the model from T1 to T4 (within session one), $z = -2.301$, $P = 0.021$, $r = 0.594$, and significant increase from T4 to T5 (from session one to session two), $z = -2.329$, $P = 0.020$, $r = 0.601$.

There was no significant change in peak-to-peak time for the model-only group, Friedman $\chi^2(4) = 4.054$, $P = 0.399$. Peak-to-peak time increased gradually toward the model in session one from T1 to T4, and then decreased in session two from T4 to T5. The model and instruction group showed a similar but more pronounced change in peak-to-peak values. This change was statistically significant, $\chi^2(4) = 12.027$, $P = 0.017$. Peak-to-peak time increased significantly toward the model from T1 to T4 (within session one), $z = -2.783$, $P = 0.005$, $r = 0.719$, followed by a small decrease in scores from T4 to T5 (from session one to session two).

No significant change in RMSE was found for both the model-only group, $\chi^2(4) = 3.947$, $P = 0.413$, and the model and instruction group, $\chi^2(4) = 0.267$, $P = 0.992$. RMSE values in both groups decreased from T1 to T4, indicating that sirens showed closer approximation to the pitch contour of the model during the first session. This was followed by an increase in values in the second session, from T4 to T5.

Interaction

To compare the change in performance between groups and answer our second research question, phonation time and intensity measures were analyzed with 2×5 repeated measures ANOVAs. The CPP, peak-to-peak time, and RMSE were analyzed with Mann-Whitney *U* tests where the two groups were compared on the difference scores between T1 and T5.

No significant interaction was found between groups for phonation time, $F(4, 25) = 1.039$, $P = 0.407$, partial $\eta^2 = 0.143$, CPP, $U = 95$, $z = -0.726$, $P = 0.468$, $r = 0.133$; peak-to-peak time, $U = 78.5$, $z = -1.410$, $P = 0.158$, $r = 0.257$; and RMSE, $U = 93$, $z = -0.809$, $P = 0.419$, $r = 0.148$. There was a significant order 4 interaction for vocal intensity, $F(4, 25) = 6.398$, $P = 0.017$, partial $\eta^2 = 0.186$. Compared with the model-only group, mean intensity in the model and instruction group showed greater improvement from T2 and T3 when instructions were first provided in session one. However, intensity values for the model and instruction group worsened slightly from T3 to T4, whereas the model-only group continued to progress during this time.

DISCUSSION

This study investigated the value of modeling and verbal instruction in the acquisition and short-term retention of a complex voice skill. Contrary to our first hypothesis, there was insufficient evidence to demonstrate that providing a model either with or without instructions was sufficient for acquisition and short-term retention over 40 trials. Although both groups showed progress in the first session, they did not successfully reproduce the model's siren performance. One day later, their performance had not changed from the baseline level, but some retention was observed as evidenced by reduced vocal intensity and higher peak-to-peak times in the model and instruction group at time 5. We also found that providing instructions with a model was more beneficial than providing a model only. An improved performance of the task was observed when the instructions were first introduced. This provides some preliminary evidence for our second hypothesis. Instructions facilitated greater initial progress in reducing vocal intensity, and the model and instruction group improved on more outcome measures compared with the model-only group although most interactions were not significant.

Acquisition and short-term retention

Although both groups showed closer approximation to the model across trials in session one, they did not successfully reproduce the model's siren performance on the outcomes measured. This shows that participants did not acquire the target skill; they did not even complete the prepractice phase according to Maas et al's[7] definition. These findings suggest that the addition of other prepractice variables or that more trials were required.

The need for other prepractice variables

Feedback may have been necessary, as the vocal siren is a complex task that was unfamiliar to participants. In contrast to simple skills, Laguna[64] found that a combination of information sources was important for complex skill acquisition. The guidance hypothesis states that an appropriate amount of feedback guides individuals to reduce error and produce more accurate movements.[65] Feedback provides vital information about the performance of each attempt relative to the target outcome.[66] Given the functional difficulty of the task, the lack of feedback information could have placed unrealistic demands on participants.[33] It may have been difficult for them to detect errors and make appropriate adjustments to match the model, leading to unsuccessful acquisition. There are two types of feedback in motor learning: knowledge of performance (KP) and knowledge of results (KR). Voice research generally favors the provision of KP when the task is unfamiliar and low-frequency KR later in the learning process.[4,7,49,67] However, there is evidence that frequent feedback can facilitate increased performance and learning for complex motor skills such as the siren.[68] More research is required to determine optimal feedback schedules for complex voice tasks during the acquisition phase of therapy.

It is also possible that the lack of perceptual training was detrimental to acquisition. If participants could accurately identify the salient features of their siren attempts, they would receive intrinsic performance feedback. However, participants' perceptual judgments may be inaccurate and inconsistent, especially considering their inexperience in voice treatment generally.[69] As all participants were vocally healthy, changes in voice characteristics were likely to be subtle and difficult to identify.[70] Perceptual training may help individuals to be more aware of sensations associated with accurate production, including what they hear and feel.[2] Research shows that such training can improve the reliability of perceptual judgments.[71] In this way, perceptual training could have facilitated the development of an internal error-detection mechanism, which may have led to acquisition of the target movement.

Number of trials

The key features of prepractice include understanding the task, establishing an internal reference of correctness, and being able to perform the task under optimal conditions.[7] No participants met these prepractice criteria, and hence they also did not achieve acquisition or short-term retention within 40 attempts. It may be that more trials were needed to achieve the goals of prepractice. When participants produce more trials, they gain more opportunities to establish, strengthen, and stabilize the movement's GMP.[72] Although there is no prepractice research regarding number of trials, a large practice amount is generally recommended to ensure motor learning.[7] Among motor learning studies, the number of practice trials ranges widely and appears to be determined arbitrarily. For example, Steinhauer and Grayhack[4] used 40 trials for acquisition of a vowel nasalization task, whereas Hodges and Franks[73] had 60 trials over 2 days for a bimanual task. As these studies did not distinguish between prepractice and practice, it is unclear if a proportion of the stated trials should be considered prepractice. Further motor learning research, particularly in SLP, should consider the value of clearly distinguishing prepractice and practice as two separate phases of acquisition in the therapeutic context.

Motor learning definitions for voice tasks

As previously described, the acquisition criterion in motor speech studies is usually set at 80% accuracy over two to three sessions or blocks.[10–12] Within acquisition, prepractice criteria include being able to perform the task accurately under optimal conditions.[7] In this study, we considered siren productions to be accurate if they matched the model on the target outcome measures. On this basis, participants did not achieve prepractice or acquire the task.

However, these criteria may be unrealistic for a complex vocal skill like the siren. A complex task requires simultaneous control and coordination of multiple parameters. It also has high functional difficulty when presented as a novel task, requiring more cognitive work during acquisition.[33] It may be unlikely that participants with no vocal training would be able to reproduce this complex siren task at the standard set by an expert voice clinician, within a few blocks of trials. Furthermore, it can be difficult to definitively classify accurate and inaccurate productions for vocal tasks. Clients may be more likely to show gradual progress over trials. As all participants approached the expert model's production in session one, approximation of the target task could be a more reasonable benchmark to illustrate acquisition. It may be that for voice, acquisition consists of prepractice as well as practice to refine the task and approach accurate performance. Once participants produce the task correctly, further practice to replicate and automatize correct productions will then aid retention and transfer. Clearly, more research is necessary to appropriately define these important concepts in the context of voice.

The role of instruction

Verbal instruction had a positive effect on acquisition, with the model and instruction group improving on a greater number of outcome measures. For vocal intensity, instructions facilitated faster progress during the second block of trials when they were first introduced. We note that mean vocal intensity in the model and instruction group worsened slightly in the third block, suggesting that continued instruction may have had adverse effects on intensity performance. It may be that participants attended to other components of the provided instructions during this time.

As they focused on changing other siren features (eg, voice quality), vocal intensity performance was compromised. This relationship between selective attention to parts, complexity of task, and relative performance has previously been reported in speech motor learning.[74] Nevertheless, our findings still suggest that verbal instruction was beneficial to the overall acquisition of a complex voice task.

Positive effects of instruction

Our positive results contradict other motor learning studies demonstrating the adverse effects of instruction on acquisition.[27,29] It has been suggested that explicit instruction causes learners to attend to limited task features and prevents them from actively testing various movement options, which can have a negative effect on learning.[28] However, restricting trial and error may actually be important in voice therapy, as producing incorrect versions of the target task may cause vocal hyperadduction and tissue damage to the vocal folds.[75]

The facilitatory effect of instruction in this study could be due to a number of factors. Greater cognitive processing is involved in gaining information from verbal instructions compared with visual demonstrations.[76] In this study, the mental work of processing the instructions may have assisted performance by helping participants better understand the task and recognize accurate productions. Second, verbal instruction provided another source of information for the performance of a novel, complex task. As functional difficulty was high under these conditions, the *challenge point framework* states that providing more information will reduce processing demands, moving participants closer to their optimal challenge points and facilitating performance.[33] Third, as participants were unfamiliar with the characteristics of vocal tasks, instruction could have also helped them to perceive important features of the model. For example, the instructions provided helped them to know that the siren task started on an *ng* sound instead of another nasal sound. Some participants who did not receive instructions used an *m* sound when trying to match the model. Other studies also support the use of verbal cues to direct attention to important model features.[9,77]

Type of instruction

These findings call for closer examination of the instructions provided. The instructions predominantly directed attention to movement effects (eg, "Gradually decrease and increase your pitch"; "Your voice should be clear and easy to produce"), rather than describing the actual movements of the vocal mechanism. This provides another explanation for the beneficial effects of instruction in this study, as external-focus instruction has consistently resulted in better acquisition.[25] In motor learning research, external-focus instruction refers to movement effects on the environment (eg, movement of the ball or racket), whereas internal-focus instruction refers to movement of body parts. However, this distinction is less clear for voice skills, as voice task production does not act upon any external instruments. This is in contrast to other motor skills such as ball sports. Maas et al[7] also found similar issues in applying attentional focus research from nonspeech domains to speech production. Hence, further research exploring the distinction between internal-focus and external-focus instruction in the context of voice and speech tasks could be beneficial.

Paradoxical CPP findings

Interestingly both groups showed a significant decrease in CPP over trials on the first day, and then a significant increase at short-term retention. Lower CPP values are associated with greater aperiodicity and poorer voice quality.[78] The CPP results hence imply that performance worsened and then improved, contradicting the results from other measures. The low CPP value of the expert model (as seen in Figure 3) also complicates the interpretation of this measure.

There are two possible explanations for these findings. First, changes in CPP may be related to changes in vocal intensity. The vocal siren should be produced at a quiet volume. As participants decreased their volume over trials, there may have been reduced vocal fold closure resulting in an increase in signal noise, in turn lowering CPP values. Joscelyne-May et al[49] found a similar relationship between intensity and CPP while exploring the effect of feedback type on the acquisition of vocal sirens. Recently, Awan, Giovinco, and Owens[79] established that intensity does have significant effect on CPP measures, with higher intensity resulting in higher CPP. Second, given that the siren is produced on an *ng* sound, the nasal properties of the siren could have affected CPP values.[80] As participants aimed to match the model, there may have been increased nasality in siren production over trials. This would have caused a greater damping effect due to the increased surface area of the nasal cavity and passages.[81] Damping spreads wave energy across a broader frequency range, resulting in a greater bandwidth and smaller spectral peaks.[82] Therefore, an increase in nasality would have reduced CPP.

Nonsignificant RMSE results

Whereas other outcome measures revealed significant results, there was no significant change in RMSE scores. To our knowledge, although RMSE has been used in several motor learning studies,[83,84] no study has used this to quantify vocal change over time. Because RMSE accounts for the entire siren, it may be that this global measure was not sensitive enough to detect changes in performance. Voice errors such as pitch breaks could have also affected RMSE scores. Nevertheless, as RMSE trends were similar to other outcome measures in this study, RMSE is a viable whole-task measure that should be explored further in voice research, especially considering that many voice measures only assess specific task components rather than overall performance.

Limitations

The findings of this study should be considered in view of some limitations. First, all participants were young, female SLP students with healthy voices. This demographic was not representative of the general population, and it also limits the generalizability of results to people with voice disorders. Second, the statistically significant changes in outcome measures may not be perceptually significant. For example, a 3 dB change in vocal intensity is typically considered the minimum difference that can be audibly detected.[85] In this study, changes in mean intensity were 2.15 dB and 2.95 dB for the model-only and for the model and instruction group, respectively. It is not known if perceptual judgment would have identified any changes in vocal intensity or other outcome measures. This is important as voice is usually evaluated based on how an individual sounds, with perceptual judgment recognized as a key component of a valid voice assessment.[86,87] Third, although this study focused on a complex skill, the application of these results to simple skills is premature. There is evidence that modeling and instruction affects simple and complex tasks in different ways.[32] As voice therapy can involve both simple and complex skills, a follow-up to this study could examine modeling and instruction in the acquisition and retention of a simple vocal skill. It is important to note that motor learning is task specific and facilitated by learner's perception of the task being meaningful. Although the siren used in this research is novel, complex, and clinically useful, it may not have been meaningful to the learners. Finally, modeling and instruction were defined generally in this paper, as the aim was to reproduce the models and instructions a client would typically receive in voice therapy or online voice training. The study did not examine the effects of various types of models and instructions, such as video self-models, learner models, external-focus instruction, and internal-focus instruction.[8] These concepts have not yet been explored in voice motor learning. Further research in this area may provide insight into the effectiveness of specific types of prepractice variables in voice.

Implications

This research has significant implications for both traditional and technology-assisted voice therapy and training. Verbal instruction is helpful when clients are acquiring a voice task. Voice clinicians should therefore provide both task demonstrations and instructions during the prepractice phase. However, as these elements were not sufficient for acquisition or short-term retention, other prepractice components such as feedback and perceptual training should be considered.

This study also suggests that exclusive use of noninteractive computer programs and mobile applications may not be useful for acquisition and facilitating accurate production of a complex vocal task. This is significant in the context of vocal impairment, as clients risk phonotrauma if they unknowingly practice hyperfunctional vocal behaviors instead of the target voice task.[75]

Interactive computer programs and mobile applications may be viable options for facilitating effective and efficient acquisition, as these tools can measure and monitor prepractice attempts. In addition to providing feedback on performance, interactive tools may also be able to detect undesirable voice characteristics to ensure vocal safety. Recently, with the aim of augmenting LSVT Companion capabilities, a pilot study examined the automatic assessment of sustained vowel phonations using a statistical framework.[88] It was shown that an algorithm could objectively classify voice samples as acceptable or unacceptable to 90% accuracy, independent of expert clinicians. This research highlights the potential of technology-assisted tools in future voice service delivery. However, before such devices are developed and rigorously tested, voice computer programs and mobile applications should be used as an adjunct to face-to-face therapy, rather than a substitute.

Although this study contributes to the voice motor learning literature, further investigation is clearly required to better understand the roles of other prepractice elements and specific types of models and instruction. These results should be replicated with a range of voice therapy techniques and vocally disordered populations, so as to establish their ecological validity. Voice therapy research is relatively scarce compared to other fields in SLP such as language and literacy.[89] We envisage that greater research interest in voice motor learning will allow voice clinicians to provide more confident therapeutic recommendations, ultimately improving the quality of voice therapy.

Acknowledgments

This project was supported by funding from The University of Sydney Dr Liang Voice Program.

REFERENCES

1. Roy N, Merrill RM, Gray SD, et al. Voice disorders in the general population: prevalence, risk factors, and occupational impact. *Laryngoscope*. 2005;115:1988–1995.
2. McIlwaine A, Madill C, McCabe P. Voice therapy prepractice and the principles of motor learning. *ACQ Knowl Speech Lang Hear*. 2010;12:29–32.
3. Ma EPM, Yiu GKY, Yiu EML. The effects of self-controlled feedback on learning of a "relaxed phonation" task. *J Voice*. 2013;27:723–728.
4. Steinhauer K, Grayhack JP. The role of knowledge of results in performance and learning of a voice motor task. *J Voice*. 2000;14:137–145.
5. Wong AYH, Ma EPM, Yiu EML. Effects of practice variability on learning of relaxed phonation in vocally hyperfunctional speakers. *J Voice*. 2011;25:e103–e113.
6. Schmidt RA, Lee TD. *Motor Control and Learning: A Behavioral Emphasis*. Champaign, IL: Human Kinetics; 2011.
7. Maas E, Robin DA, Austermann Hula SN, et al. Principles of motor learning in treatment of motor speech disorders. *Am J Speech Lang Pathol*. 2008;17:277–298.
8. Hodges NJ, Franks IM. Modelling coaching practice: the role of instruction and demonstration. *J Sports Sci*. 2002;20:793–811.
9. Hodges NJ, Franks IM. Instructions, demonstrations and the learning process. In: Williams AM, Hodges NJ, eds. *Skill Acquisition in Sport: Research, Theory and Practice*. London: Routledge; 2004:145–174.
10. Ballard KJ, Robin DA, McCabe P, et al. A treatment for dysprosody in childhood apraxia of speech. *J Speech Lang Hear Res*. 2010;53:1227–1245.
11. Ruscello DM. *Treating Articulation and Phonological Disorders in Children*. St. Louis, MO: Mosby Elsevier; 2008.
12. Thomas DC, McCabe P, Ballard KJ. Rapid syllable transitions (ReST) treatment for childhood apraxia of speech: the effect of lower dose-frequency. *J Commun Disord*. 2014;51:29–42.
13. Bislick LP, Weir PC, Spencer K, et al. Do principles of motor learning enhance retention and transfer of speech skills? A systematic review. *Aphasiology*. 2012;26:709–728.
14. Raab M, Masters RSW, Maxwell J, et al. Discovery learning in sports: implicit or explicit processes? *Int J Sport Exerc Psychol*. 2009;7:413–430.
15. Ramig LO, Verdolini K. Treatment efficacy: voice disorders. *J Speech Lang Hear Res*. 1998;41:S101–S116.
16. Schmidt RA. Motor schema theory after 27 years: reflections and implications for a new theory. *Res Q Exerc Sport*. 2003;74:366–375.
17. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall; 1986.
18. Ram N, Riggs SM, Skaling S, et al. A comparison of modelling and imagery in the acquisition and retention of motor skills. *J Sports Sci*. 2007;25:587–597.
19. Hodges NJ, Williams AM, Hayes SJ, et al. What is modelled during observational learning? *J Sports Sci*. 2007;25:531–545.

20. Shea CH, Wright DL, Wulf G, et al. Physical and observational practice afford unique learning opportunities. *J Mot Behav*. 2000;32:27–36.
21. Weeks DL, Anderson LP. The interaction of observational learning with overt practice: effects on motor skill learning. *Acta Psychol (Amst)*. 2000;104:259–271.
22. Molenberghs P, Cunnington R, Mattingley JB. Is the mirror neuron system involved in imitation? A short review and meta-analysis. *Neurosci Biobehav Rev*. 2009;33:975–980.
23. Lago-Rodriguez A, Lopez-Alonso V, Fernández-del-Olmo M. Mirror neuron system and observational learning: behavioral and neurophysiological evidence. *Behav Brain Res*. 2013;248:104–113.
24. Swinnen SP. Information feedback for motor skill learning: a review. In: Zelaznik HN, ed. *Advances in Motor Learning and Control Champaign*. Human Kinetics; 1996:37–66.
25. Wulf G. Attentional focus and motor learning: a review of 15 years. *Int Rev Sport Exerc Psychol*. 2013;16:77–104.
26. Masters RSW, Poolton J. Advances in implicit motor learning. In: Hodges NJ, Williams AM, eds. *Skill Acquisition in Sport: Research, Theory and Practice*. 2nd ed. London: Routledge; 2012:59–75.
27. Wulf G, Höß M, Prinz W. Instructions for motor learning: differential effects of internal versus external focus of attention. *J Mot Behav*. 1998;30:169–179.
28. Hodges NJ, Lee TD. The role of augmented information prior to learning a bimanual visual-motor coordination task: do instructions of the movement pattern facilitate learning relative to discovery learning? *Brit J Psychol*. 1999;90:389–403.
29. Wulf G, Weigelt C. Instructions about physical principles in learning a complex motor skill: to tell or not to tell. *Res Q Exerc Sport*. 1997;68:362–367.
30. Lawrie R, Madill C, Heard R, et al. A comparison of the clinical behaviors of students and experienced speech pathologists in the prepractice phase of voice therapy: a preliminary investigation. Voice Foundation 40th Annual Symposium (Care of the Professional Voice); 2011; Philadelphia, USA.
31. Fontana FE, Mazzardo O, Furtado JO, et al. Whole and part practice: a meta-analysis. *Percept Mot Skills*. 2009;109:517–530.
32. Wulf G, Shea CH. Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychon Bull Rev*. 2002;9:185–211.
33. Guadagnoli MA, Lee TD. Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. *J Motor Behav*. 2004;36:212–224.
34. Akay M, Tamura T. Global healthcare: advances and challenges. *P IEEE*. 2015;103:147–149.
35. van Leer E. Using popular mobile devices in voice therapy. *Perspect Voice Voice Disord*. 2013;23:82–87.
36. VoixTek. VoixTek voice training; 2013.
37. On Cue Productions. Vocal ease; 2014.
38. Vocal Splendor Studios. The Hum—vocal exercises, part I; 2012.
39. National Center for Voice & Speech. Vocal straw exercise. [Video file]; 2010. Available at: https://www.youtube.com/watch?v=0xYDvwvmBIM. Accessed June 17, 2017.
40. Halpern AE, McFarland DH, Ramig LO, et al. Innovative technology for the assisted delivery of intensive voice treatment (LSVT®LOUD) for Parkinson disease. *Am J Speech Lang Pathol*. 2012;21:354–367.
41. King SN, Davis L, Lehman JJ, et al. A model for treating voice disorders in school-age children within a video gaming environment. *J Voice*. 2012;26:656–663.
42. Mint Leaf Software Inc. Sonneta Voice Monitor 2.0.1; 2014.
43. Cantovation Ltd. Sing & See Professional USYD Speech Science Special Build. 1.5.7 ed; 2013.
44. Brady G. Voice Analyst 1.2; 2014.
45. OperaVOX Ltd. OperaVOX 1.1; 2013.
46. Russell R, Pemberton C, Oates J. Voice Care for Teachers DVD: A prevention DVD. 2007.
47. Catona G. Voice Builder; 2014.
48. Musically Speaking. Virtual Voice Lessons 1.3; 2013.
49. Joscelyne-May C, McCabe P, Thorpe W, et al. The effect of clinician feedback type in the acquisition of a vocal siren. *Folia Phoniatr Logop*. 2015;67:57–67.
50. American Speech-Language-Hearing Association. Guidelines for manual pure-tone threshold audiometry; 2005. Available at: http://www.asha.org/policy/GL2005-00014/.
51. Simberg S, Sala E, Laine A, et al. A fast and easy screening method for voice disorders among teacher students. *Logoped Phoniatr Vocol*. 2001;26:10–16.
52. Awan SN, Morrow DL. Videostroboscopic characteristics of young adult female smokers vs. nonsmokers. *J Voice*. 2007;21:211–223.
53. Ihre E, Zetterstrom O, Ihre E, et al. Voice problems as side effects of inhaled corticosteroids in asthma patients—a prevalence study. *J Voice*. 2004;18:403–414.
54. Harris T, Harris S, Rubin JS, et al. *The Voice Clinic Handbook*. 1 ed. United Kingdom: Wiley; 1997.
55. Stemple JC, Lee L, Damico B, et al. Efficacy of vocal function exercises as a method of improving voice production. *J Voice*. 1994;8:271–278.
56. Sapir S, Ramig LO, Fox CM. Intensive voice treatment in Parkinson's disease: Lee Silverman voice treatment. *Expert Rev Neurother*. 2011;11:815–830.
57. NCH. WavePad Audio Editing Software [computer program]: NCH Sofware. 2011. Available at: http://www.nch.com.au/wavepad/index.html.
58. Boersma P, Weenink D. Praat:doing phonetics by computer [computer program] (Version 5.3). 2014. Available at: http://www.praat.org/.
59. Maslan J, Leng X, Rees C, et al. Maximum phonation time in healthy older adults. *J Voice*. 2011;25:709–713.
60. Watts CR, Awan SN. Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts. *J Speech Lang Hear Res*. 2011;54:1525–1537.
61. Hillenbrand J. SpeechTool. 2011. http://homepages.wmich.edu/~hillenbr/.
62. Kelley K, Lai K. Accuracy in parameter estimation for the root mean square error of approximation: sample size planning for narrow confidence intervals. *Multivariate Behav Res*. 2011;46:1–32.
63. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
64. Laguna PL. Task complexity and sources of task-related information during the observational learning process. *J Sports Sci*. 2008;26:1097–1113.
65. Schmidt RA. Frequent augmented feedback can degrade learning: evidence and interpretations. In: Requin J, Stelmach GE, eds. *Tutorials in Motor Neuroscience*. The Netherlands: Kluwer Academic; 1991:59–75.
66. Wulf G, Shea CH, Lewthwaite R. Motor skill learning and performance: a review of influential factors. *Med Educ*. 2010;44:75–84.
67. Ferrand CT. Effects of practice with and without knowledge of results on jitter and shimmer levels in normally speaking women. *J Voice*. 1995;9:419–423.
68. Wulf G, Shea CH, Matschiner S. Frequent feedback enhances complex motor skill learning. *J Mot Behav*. 1998;30:180–192.
69. Kreiman J, Gerratt BR, Kempster GB, et al. Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *J Speech Hear Res*. 1993;36:21–40.
70. Rabinov CR, Kreiman J, Gerratt BR, et al. Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *J Speech Hear Res*. 1995;38:26–32.
71. Eadie TL, Baylor CR. The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *J Voice*. 2006;20:527–544.
72. Whitacre CA, Shea CH. Performance and learning of generalized motor programs: relative (GMP) and absolute (parameter) errors. *J Mot Behav*. 2000;32:163–175.
73. Hodges NJ, Franks IM. Attention focusing instructions and coordination bias: implications for learning a novel bimanual task. *Hum Mov Sci*. 2000;19:843–867.
74. McCabe P, Macdonald-D'Silva AG, van Rees LJ, et al. Orthographically sensitive treatment for dysprosody in children with childhood apraxia of speech using ReST intervention. *Dev Neurorehabil* 2014;17:137–145.
75. Franco RA, Andrus JG. Common diagnoses and treatments in professional voice users. *Otolaryngol Clin North Am*. 2007;40:1025–1061.
76. Craik FIM, Lockhart RS. Levels of processing: a framework for memory research. *J Verbal Learning Verbal Behav*. 1972;11:671–684.

77. Janelle CM, Champenoy JD, Coombes SA, et al. Mechanisms of attentional cueing during observational learning to facilitate motor skill acquisition. *J Sports Sci*. 2003;21:825–838.
78. Heman-Ackah YD, Heuer RJ, Michael DD, et al. Cepstral peak prominence: a more reliable measure of dysphonia. *Ann Otol Rhinol Laryngol*. 2003;112:324–333.
79. Awan SN, Giovinco A, Owens J. Effects of vocal intensity and vowel type on cepstral analysis of voice. *J Voice*. 2012;26:670.e615–670.e620.
80. Eastwood C, Madill C, McCabe P, et al. Vocal clarity of female speech pathology students. Speech Pathology Australia Annual National Conference; 2012; Hobart, Australia.
81. Schwartz MF. The acoustics of normal and nasal vowel production. *Cleft Palate J*. 1968;5:125–140.
82. Pruthi T, Espy-Wilson CY. Acoustic parameters for automatic detection of nasal manner. *Speech Commun*. 2004;43:225–239.
83. Sullivan KJ, Kantak SS, Burtner PA. Motor learning in children: feedback effects on skill acquisition. *Phys Ther*. 2008;88:720–732.
84. Wulf G, McNevin N, Shea CH. The automaticity of complex motor skill learning as a function of attentional focus. *Q J Exp Psychol*. 2001;54:1143–1154.
85. Sum KS, Pan J. Subjective evaluation of reverberation times of sound fields with non-exponential decays. *Acta Acust United Acust* 2006;92:583–592.
86. Bhuta T, Patrick L, Garnett JD. Perceptual evaluation of voice quality and its correlation with acoustic measurements. *J Voice*. 2004;18:299–304.
87. Carding P, Carlson E, Epstein R, et al. Formal perceptual evaluation of voice quality in the United Kingdom. *Logopedics Phonatrics Vocology*. 2000;25:133–138.
88. Tsanas A, Little MA, Fox C, et al. Objective automatic assessment of rehabilitative speech treatment in Parkinson's disease. *IEEE Trans Neural Syst Rehabil Eng* 2014;22:181–190.
89. Manoharan D, Togher L, Power E, et al. *The Reported Methodological Quality of Speech-Language Pathology Treatment Research*. Sydney, Australia: The University of Sydney; 2014.