Practical elicitation methods for the voice range profile

Ann Kolker

Follow this and additional works at: https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations

Recommended Citation
https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations/4460

This Dissertation/Thesis is brought to you for free and open access by the Graduate Research & Artistry at Huskie Commons. It has been accepted for inclusion in Graduate Research Theses & Dissertations by an authorized administrator of Huskie Commons. For more information, please contact jschumacher@niu.edu.
ABSTRACT

PRACTICAL ELICITATION METHODS FOR THE VOICE RANGE PROFILE

Ann Kolker, M.A.
School of Allied Health and Communicative Disorders
Northern Illinois University, 2017
Miriam van Mersbergen & In-sop Kim, Co-Directors

Voice production is a complex process involving the coordination of various anatomical structures and physiologic systems. The Voice Range Profile (VRP) is an acoustic measure for evaluating voice production that provides information regarding how these structures and systems function together. The present study examined shortened protocols for VRP elicitation to reduce elicitation time and allow for more widespread use of the VRP. Twenty-four singers completed a full VRP based on accepted methods from the literature, and the same participants returned within one to three weeks to complete the proposed short VRP protocol. Results indicate that the new short protocol allowed participants to generate larger VRPs with greater maximum intensities and lower minimum intensities. However, the short protocol also presented challenges for participants due to its dynamic sampling method and for elicitors due to increased cognitive load. Future research directions include generating a new rigid sampling method, using a larger sample size that includes non-singers, including additional measures, examining elicitor cognitive load or comfort, and examining participant comfort.
PRACTICAL ELICITATION METHODS FOR THE VOICE RANGE PROFILE

BY

ANN KOLKER
©2017 Ann Kolker

A THESIS SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
MASTER OF ARTS

SCHOOL OF ALLIED HEALTH AND COMMUNICATIVE DISORDERS

Thesis Co-Directors:
Miriam van Mersbergen, Ph.D., CCC-SLP
In-sop Kim, Ph.D., CFY-SLP
ACKNOWLEDGEMENTS

This study was completed with the help of many people. First, I would like to thank my mentor, Miriam van Mersbergen, for her guidance and support throughout the entire project. Without her expertise and guidance, this research project would not have been completed. Second, I would like to thank Matthew Rychel for his assistance with data analysis and his software expertise. Furthermore, I would like to thank the committee members, Dr. Allison Gladfelter, Dr. In-sop Kim, Dr. Orna Arania, and Dr. Mary Lynn Doherty, for their guidance in making this study more thorough and complete. In addition, I would like to thank all of the locations and people who allowed me to solicit participants and/or provided space for me to elicit VRPs: the Northern Illinois University Music Department, Saints Peter and Paul Catholic Church Music Department, Saint Irene Church, Saint Xavier University Music Department, Batavia High School Music Department, North Central College, and Dr. Jan Bickel. Finally, I would like to thank everyone who participated in this study.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td>viii</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>1</td>
</tr>
<tr>
<td>Literature Review</td>
<td>3</td>
</tr>
<tr>
<td>History</td>
<td>3</td>
</tr>
<tr>
<td>Factors That Affect the VRP</td>
<td>6</td>
</tr>
<tr>
<td>Production Methods</td>
<td>10</td>
</tr>
<tr>
<td>Analyzing VRP Results</td>
<td>13</td>
</tr>
<tr>
<td>Clinical Utility</td>
<td>15</td>
</tr>
<tr>
<td>Considerations for the Current Study</td>
<td>18</td>
</tr>
<tr>
<td>Purpose</td>
<td>20</td>
</tr>
<tr>
<td>2. METHODS</td>
<td>23</td>
</tr>
<tr>
<td>Participants</td>
<td>23</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>24</td>
</tr>
<tr>
<td>Procedures</td>
<td>24</td>
</tr>
<tr>
<td>Full Protocol</td>
<td>24</td>
</tr>
<tr>
<td>Shortened Protocols</td>
<td>28</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Data collection</td>
<td>29</td>
</tr>
<tr>
<td>Analysis</td>
<td>29</td>
</tr>
<tr>
<td>Time Comparisons</td>
<td>30</td>
</tr>
<tr>
<td>Frequency Differences</td>
<td>30</td>
</tr>
<tr>
<td>Intensity Differences</td>
<td>31</td>
</tr>
<tr>
<td>Criteria for Acceptance</td>
<td>33</td>
</tr>
<tr>
<td>3. RESULTS</td>
<td>34</td>
</tr>
<tr>
<td>Time Comparisons</td>
<td>34</td>
</tr>
<tr>
<td>Full vs. Short Comparison</td>
<td>36</td>
</tr>
<tr>
<td>Male Frequency Differences</td>
<td>36</td>
</tr>
<tr>
<td>Female Frequency Differences</td>
<td>37</td>
</tr>
<tr>
<td>Male Intensity Differences</td>
<td>38</td>
</tr>
<tr>
<td>Female Intensity Differences</td>
<td>42</td>
</tr>
<tr>
<td>Sparse vs. Short Comparison</td>
<td>46</td>
</tr>
<tr>
<td>Frequency Differences for Males and Females</td>
<td>46</td>
</tr>
<tr>
<td>Male Intensity Differences</td>
<td>46</td>
</tr>
<tr>
<td>Female Intensity Differences</td>
<td>49</td>
</tr>
<tr>
<td>Sparse vs. CEGA Comparison</td>
<td>54</td>
</tr>
<tr>
<td>Frequency Differences for Males and Females</td>
<td>54</td>
</tr>
<tr>
<td>Male Intensity Differences</td>
<td>54</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Female Intensity Differences</td>
<td>57</td>
</tr>
<tr>
<td>4. DISCUSSION</td>
<td>62</td>
</tr>
<tr>
<td>Elicitation Considerations</td>
<td>63</td>
</tr>
<tr>
<td>Limitations to the Research</td>
<td>67</td>
</tr>
<tr>
<td>Future Directions</td>
<td>69</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>72</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>75</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.</td>
<td>Male and female average elicitation time, number of sampling points</td>
</tr>
<tr>
<td></td>
<td>(minimum plus maximum curve), and time per sampling point</td>
</tr>
<tr>
<td>2.</td>
<td>Male and female two-sample equal-variance t tests comparing full and</td>
</tr>
<tr>
<td></td>
<td>short elicitation time</td>
</tr>
<tr>
<td>3.</td>
<td>Average range differences at the low and high ends of the range between</td>
</tr>
<tr>
<td></td>
<td>short and full protocols for males and females</td>
</tr>
<tr>
<td>4.</td>
<td>Average frequency range differences by number of semitones between the</td>
</tr>
<tr>
<td></td>
<td>short and full protocol for male and females</td>
</tr>
<tr>
<td>5.</td>
<td>Male and female two-sample equal variance t tests comparing full and short</td>
</tr>
<tr>
<td></td>
<td>protocols on the minimum and maximum curves</td>
</tr>
<tr>
<td>6.</td>
<td>Preliminary measurements of ambient noise in various environments using</td>
</tr>
<tr>
<td></td>
<td>dB C and dB A weighting</td>
</tr>
<tr>
<td>7.</td>
<td>Average ambient noise and ambient noise range in elicitation environments</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Sample Voice Range Profile</td>
<td>4</td>
</tr>
<tr>
<td>2.</td>
<td>Descending, ascending, descending arpeggio for semitone range elicitation</td>
<td>25</td>
</tr>
<tr>
<td>3.</td>
<td>Average male full VRP graph compared to average male short VRP graph</td>
<td>39</td>
</tr>
<tr>
<td>4.</td>
<td>Average male full versus short SPL difference scores across frequency range</td>
<td>42</td>
</tr>
<tr>
<td>5.</td>
<td>Average female full VRP graph compared to average female short VRP graph</td>
<td>43</td>
</tr>
<tr>
<td>6.</td>
<td>Average female full versus short SPL difference scores across frequency range</td>
<td>45</td>
</tr>
<tr>
<td>7.</td>
<td>Average male sparse VRP graph compared to average male short VRP graph</td>
<td>47</td>
</tr>
<tr>
<td>8.</td>
<td>Average male sparse versus short SPL difference scores across frequency range</td>
<td>49</td>
</tr>
<tr>
<td>9.</td>
<td>Average female sparse VRP graph compared to average female short VRP graph</td>
<td>50</td>
</tr>
<tr>
<td>10.</td>
<td>Average female sparse versus short SPL difference scores across frequency range</td>
<td>53</td>
</tr>
<tr>
<td>11.</td>
<td>Average male sparse VRP graph compared to average male CEGA VRP graph</td>
<td>55</td>
</tr>
<tr>
<td>12.</td>
<td>Average male sparse versus CEGA SPL difference scores across frequency range</td>
<td>57</td>
</tr>
<tr>
<td>13.</td>
<td>Average female sparse VRP graph compared to average female CEGA VRP graph</td>
<td>58</td>
</tr>
<tr>
<td>14.</td>
<td>Average female sparse versus CEGA SPL difference scores across frequency range</td>
<td>60</td>
</tr>
<tr>
<td>15.</td>
<td>Background noise correction</td>
<td>77</td>
</tr>
<tr>
<td>16.</td>
<td>A graph illustrating the difference between A and C sound level meter weighting</td>
<td>78</td>
</tr>
</tbody>
</table>
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td>SUMMARY OF PRELIMINARY AND ACTUAL AMBIENT NOISE MEASUREMENTS</td>
<td>75</td>
</tr>
<tr>
<td>B.</td>
<td>VHI-30</td>
<td>81</td>
</tr>
<tr>
<td>C.</td>
<td>PARTICIPANT INTAKE</td>
<td>83</td>
</tr>
<tr>
<td>D.</td>
<td>ELICITATION INSTRUCTIONS</td>
<td>85</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

Background

Voice production is a complex process involving the coordination of various anatomical structures and physiologic systems (Titze, 2000, p. 211-280). An individual’s vocal capabilities are influenced by his or her use of the respiratory system, laryngeal mechanism, articulators, and resonators. Therefore, assessment of voice production in its entirety requires multiple measures performed by experienced professionals. Voice researchers often have extended time periods to assess voice function; however, speech-language pathologists, voice teachers, choir directors, and other voice professionals may not have this same luxury. Researchers, speech-language pathologists and other voice professionals need an efficient and global method for objectively analyzing vocal function and progress.

Perceptual measures may be used to determine the quality of a voice based on how the voice sounds to an experienced listener (Mehta & Hillman, 2008). However, these measures are subjective, and they cannot be compared to norms without extensive training (Kempster, Gerratt, Verdolini Abbott, Barkmeier-Kraemer, & Hillman, 2009). On the other hand, acoustic voice parameters can be compared among speakers or singers to determine the quality or capabilities of the voice. Some acoustic measures analyze the voice and compare recorded values to normative data, such as frequency range, fundamental frequency, jitter, and shimmer. Although these measures can be compared to normative data, they do not provide a comprehensive picture of
voice function. In addition, these measures may require expensive instrumentation. Aerodynamic measures are also used to evaluate voice function (Mehta & Hillman, 2008). Aerodynamic measures typically involve estimates of phonation threshold pressure, subglottal pressure, and glottal airflow rates. Although these measures are not invasive, they also require expensive instrumentation and only evaluate a specific component of voice production.

The Voice Range Profile (VRP) is a practical measure that examines voice capability and does not necessarily require expensive instruments. The VRP samples maximum and minimum intensities at specific frequencies within the client’s frequency range. This acoustic measure generates a graph that reveals a picture of unified voice function. In addition, this measure reveals a client’s vocal capabilities, or the physiologic voice limits for this individual based on how the individual uses the voice (DeJonckere, van Wijck, & Speyer, 2003; Sulter, Schutte, & Miller, 1995; Titze & Hunter, 2011). The Voice Range Profile provides information not only about voice capabilities but also about progress following intervention, voice maturation, and how an individual’s voice compares to normative data based on individuals in the same population (Sulter, Wit, Schutte, & Miller, 1994). Despite the practical applications of this measure, standard protocols require updates for technological advances, and traditional elicitation methods can be lengthy and fatiguing for the client.
Literature Review

History

The Voice Range Profile (VRP) is an acoustic voice measure that plots intensity in sound pressure level (SPL) versus frequency in semitones or Hertz (Hz) (Sanchez, Oates, Dacakis, & Holmberg, 2014; Titze, 1992). A VRP is a physiologic measure that generates a visual representation of a person’s maximum voice capacity or minimum and maximum extremes of phonation (Pabon, Ternström, & Lamarche, 2011; Sanchez et al., 2014). The graphic output consists of two curves. The upper contour represents the loudest possible phonation at points across a person’s entire frequency range (Hallin, Frost, Holmberg, & Sodersten, 2012). The lower curve represents the softest possible phonation across a person’s entire frequency range. Together, the curves comprise a visual representation of voice function and physiologic capabilities or limits (Coleman, 1993; Sanchez et al., 2014). Most VRPs have a characteristic shape often described as two overlapping ellipses with the second ellipse tilted upward as frequency increases (Sulter, Wit, Schutte, & Miller, 1994; Titze, 2000, p. 260-261; see Figure 1).
Figure 1: Sample Voice Range Profile. This figure is a graphic display of a VRP.

Historically, the VRP has been known as the phonetogram, stimmfeld, courbe vocale, or frequency-intensity graph (Schutte & Seidner, 1983). In 1935, Wolf and Sette devised a measure for examining singing voice power that preceded the VRP (Coleman, Mabis, & Hinson, 1977). In this study, they examined the relationship between frequency and intensity in voice production. They found that when four male singers produced their loudest phonations across their frequency ranges, intensity constantly increased as frequency increased for about two octaves of the participants’ ranges. Outside of this two-octave range, intensity increased, but the increase was not consistent. This first study included only four participants, but a later study by Wolf, Stanley, and Sette (1935) with over 50 participants revealed a consistent increase in intensity as frequency increased. Another study by Stout (1938) established the modern version of the VRP with both minimum and maximum-intensity contours. In his study, Stout (1938) also determined that vowel choice can affect the relationship between frequency and intensity on the
VRP (Coleman et al., 1977). Some vowels exhibited greater intensity increases with increasing frequency while others exhibited less pronounced intensity increases with increasing frequency (Stout, 1938). By the 1970s, the VRP had become a more popular measure of voice function (Coleman, 1993). By the 1980s researchers had devised automated computerized methods for eliciting VRPs.

The characteristic shape of the VRP can be explained by the physiologic phenomena observed during phonation. A correlation between frequency and intensity occurs, especially when the vowel /a/ is used for VRP elicitation (Titze, 2000, p. 260-261; Wolf, Stanley, & Sette, 1935). As frequency increases, so does intensity. This relationship broadly exists in both the upper curve and the lower curve. This is a result of the relationship of subglottic pressure and laryngeal resistance when the vocal folds stiffen (Titze, 2000, p. 234). Additionally, as frequency increases, more energy in the spectrum is shifted to the fundamental frequency, which results in a greater intensity. Therefore, greater intensities should be found in the upper curve at higher frequencies. However, this also means that it will be more difficult to achieve softer phonation for the lower curve as frequency increases. Therefore, intensity range is reduced at the extremes of the person’s frequency range (Titze, 1992). Intensity range is often greatest in the middle of a person’s frequency range due to the numerous ways in which to achieve frequency through muscular activation resulting in increased variability in the respiratory demands during phonation (Titze, 2000, p. 263). Due to physiologic constraints and tuning of harmonics to formants, the contours contain intensity variations that appear as ripples in the upper curve (Titze, 1992; Titze, 2000, p. 262). Lower frequencies have intensities that are lower overall because the vocal folds are laxer and cannot vibrate well with increases in lung pressure. In
addition, the resonance properties of the oral cavity do not assist in projecting lower frequencies (Titze, 2000, p. 263-264.)

**Factors That Affect the VRP**

As the Voice Range Profile grew in popularity, researchers examined the many factors that affected the graphic representation. These factors contribute to shape, frequency range, and intensity range. Understanding these factors plays an important role in the creation of norms or in comparisons of different VRPs.

**Sampling Intervals**

The sampling intervals affect VRP shape (Coleman, 1993). One can sample at every semitone in a person’s frequency range, only at 10% increments within the person’s range, or even at four semitones per octave (e.g., on semitones C, E, G, A) with additional samples at the extremes of the frequency range (Coleman, 1993; Gramming, Sundberg & Akerlund, 1991; Lamarche, Ternström & Pabon, 2008; Sulter, Schutte & Miller, 1995). While sampling at fewer semitones decreases elicitation time and therefore participant fatigue, a VRP with fewer data points will provide less information about important transitions, such as from modal to falsetto register or between formant frequencies, and will produce a less detailed shape (Coleman, 1993).

**Vowel Choice**

Another important factor is vowel choice (Coleman, 1993; Gramming, Gauffin, & Sundberg, 1986; Lamesch, Doval, & Castellengo, 2012; Titze, 2000, p. 260-261). Stout (1938) found that the vowel /a/ generated the largest intensity ranges across a person’s frequency range, and most studies have used the vowel /a/ for VRP elicitation (Coleman, Mabis, & Hinson, 1977; Gramming et al., 1986; Gramming et al., 1991; Lycke, Decoster, Ivanova, Van Hulle, & de Jong,
2012; Sanchez, Oates, Dacakis, & Holmberg, 2014; Sulter et al., 1995). However, researchers have used vowels /a/, /i/, /u/, and /o/ to elicit VRPs (Coleman, 1993; Lamesch et al., 2012). Because different vowels have different formant spacing, certain vowels may have different gains in energy at different frequencies due to tuning of harmonics with formants (Titze, 2000, p. 262). For example, a VRP elicited with the vowel /a/ will have a different shape than a VRP elicited from the same person with the vowel /i/. This different shape will reflect not only radiation characteristics of mouth opening but also the cluster of formant 1 and formant 2 observed in the vowel /i/ but not in the vowel /a/.

**Mouth Opening**

For untrained participants with no singing experience, mouth opening may also influence the vowel (Coleman, 1993). Although trained singers can increase loudness by controlling degree of mouth opening or constriction, untrained singers may change degree of mouth opening randomly. In addition, at the extremes of the participant’s frequency range, appropriate mouth opening for the desired vowel may not be maintained. If the vowel changes at the extremes, VRP shape may be affected (Coleman, 1993; Hallin et al., 2012).

**Vocal Registers**

A study by Lamesch et al. (2012) suggests that intensity characteristics of vowels can also be influenced by vocal register. Different registers depend upon the actions of the resonators and vocal folds. The three different registers are modal (chest), falsetto (head), and pulse (fry). Modal register involves thicker vocal folds and the use of more muscle, and falsetto register involves thin, stretched vocal folds with little to no muscle engagement (Lamesch et al., 2012). The researchers found that when participants used modal register, vowel choice
significantly affected the maximum-intensity curve from $C_3$ (130 Hz) to $F_4$ (349 Hz) in both males and females. Because male classical singers primarily use modal register and female classical singers use more falsetto or mixed register (blend of falsetto and modal), this finding may prove more useful when eliciting VRPs from male classical singers rather than females. Furthermore, this finding can contribute to increased VRP variability in the upper curve for both male and female untrained participants.

**Elicitation Methods**

Researchers may use discrete steady-state productions or glissando productions to elicit frequency and intensity range (Coleman, 1993). Steady-state productions involve matching a pitch and holding the pitch for a predetermined amount of time. Glissando productions involve sliding or gliding from semitone to semitone. While manual VRP elicitation requires a tone duration of 2 to 3 seconds for the elicitor to read intensity from a sound level meter, computerized VRP programs only require a tone duration of about 25 milliseconds to record frequency and intensity data. Although glissando productions may generate a larger semitone range by about two or three semitones, they are primarily used with computerized VRP recording programs. Steady-state productions are easier to use during manual elicitation, but untrained participants may have difficulty matching pitches for steady-state production (Hallin et al., 2012). Therefore, untrained participants may be more successful with computerized VRP programs that allow for glissando productions. Because mode of production can affect VRP contours, manually elicited and computer-elicited VRPs may not provide easily comparable data (Sanchez et al., 2014).
Warm-ups

In addition, more specific procedures during elicitation affect VRP shape. For example, although the exact physiological effects of warm-ups are not yet fully understood, clinicians and voice professionals agree that warming up often allows for approximation of more semitones (Coleman, 1993). Furthermore, Coleman (1993) suggests that warming up can also reassure VRP participants, whether they are singers or untrained participants.

Repeated Productions

Studies have also revealed that allowing repeated productions for each semitone generates an expanded and more physiologically representative VRP. This effect is especially noticeable in the minimum curve and for untrained participants (Coleman, 1993).

Voice Quality

Although the VRP is a physiologic measure, researchers or clinicians must choose whether to include tones produced with breathy voice quality, vibrato, glottal fry, or whistle tones (Coleman, 1993; Hallin et al., 2012). Therefore, some researchers suggest including perceptual voice quality observations as a supplement to the VRP or reporting which tone qualities were included and excluded from the final VRP (Hallin et al., 2012).

Environmental and Equipment Factors

Additional environmental and equipment factors influence the VRP. Because the VRP is an acoustic measure, room acoustics can affect intensity measurements (Coleman, 1993). When a sound-proofed booth or an anechoic chamber is not used for elicitation, rooms with ambient noise greater than the intensity of the tone produced will affect intensity measures in the lower
contour. To retain minimum-intensity data in environments with significant ambient noise, the sound level meter is often set on the dB A mode, which filters out low-frequency environmental noise (Coleman, 1993). In addition, mouth to microphone distance can be reduced from the recommended 30 centimeters to reduce the influence of ambient noise (Schutte & Seidner, 1983). However, a correction factor should be applied if mouth to microphone distance is reduced (e.g., if 30 cm distance is reduced by half to 15 cm, 6 dB should be subtracted as a correction factor to facilitate comparison among VRPs, provided there is a stable environment that complies with the inverse square law).

As is evident, many factors affect VRP shape. These factors make establishing norms and comparing VRPs difficult. Lack of standard elicitation procedures may affect VRP reliability as a tool for comparing voice production abilities within and across populations (Pabon, Ternström, & Lamarche, 2011). Although current recommendations by the European Union of Phoniatriics are often cited in research articles, this equipment and recording environment may not be practical for clinical use (Schutte & Seidner, 1983). If a practical and standard method can be widely used, VRP reliability and usefulness across various settings may increase.

Production Methods

Standard Recommendations

Despite the recommendations published by the Union of European Phoniatriics, a wide variety of elicitation methods are used throughout the literature (Pabon et al., 2011; Schutte & Seidner, 1983). The Union of European Phoniatriics suggests recording VRPs on vowels /a/, /u/, and /i/, a 30 cm mouth to microphone distance, slow dB A sound level meter settings, and a
room with “living room acoustics” that does not have “excessive damping” (Schutte & Seidner, 1983, p. 287).

Elicitation Setting

Most researchers choose to elicit VRPs in sound-treated rooms or recording studios (Gramming, Sundberg, & Akerlund, 1991; Hallin et al., 2012; Lamarche, Ternström, & Pabon, 2010; Sanchez et al., 2014; Sulter, Schutte, & Miller, 1995). In other studies, researchers recorded VRPs in anechoic chambers (Coleman et al., 1977; Gramming et al., 1986). However, these rooms are not readily available to researchers or clinicians, and the recommendation for “living room acoustics” (Schutte & Seidner, 1983, p. 287) remains poorly defined.

Mouth to Microphone Distance

Slight variations in mouth to microphone distance are recommended to compensate for reverberation or normal room acoustics (Schutte & Seidner, 1983). Many current studies either follow the 30 cm mouth to microphone distance recommendation or modify the distance and apply a correction factor so that results are comparable to VRPs elicited with the recommended 30 cm distance (Gramming et al., 1991; Hallin et al., 2012; Lamarche, Ternström, & Pabon, 2010; Lycke et al., 2012; Sulter et al., 1995). However, the close microphone to mouth distance reduces the effective reverberant field, which may create a near-field effect, producing errors in these correction factors. In addition, although adjustments in mouth to microphone distance may correct for ambient noise, one study which used a 5 cm mouth to microphone distance encountered high levels of microphone distortion above 115 dB (Sanchez et al., 2014).
**Computerized and Manual Methods**

Researchers who elicited VRPs manually first had participants record the lower contour by beginning at a comfortable middle semitone and phonating at their lowest intensity on this semitone (Gramming et al., 1991; Sulter et al., 1995). Then, participants descended from the middle semitone, obtaining minimum intensities as they descended to their minimum frequency. After obtaining the minimum intensities for the middle and low frequencies, the participants again began at the middle semitone and then ascended to gain minimum-intensity data for the higher frequencies. A similar process was repeated for the upper contour.

Computerized VRP programs have made different elicitation methods such as glissandi, shorter phonation times, and supplemental information more available (DeJonckere, van Wijck, & Speyer, 2003; Hallin et al., 2012; Lamarche, Ternström, & Pabon, 2010; Pabon et al., 2011; Sanchez et al., 2014). Computerized VRP elicitation often consists of subjects using a glissando method to expand intensity range and “fill in” as much of the VRP screen as possible (Sanchez et al., 2014). Manual elicitation often involves discrete pitch matching, often on semitones C, E, G, and A, in each octave of a participant’s range with additional semitones at the extremes of the frequency range (Gramming, Sundberg, & Akerland, 1991; Lamarche et al., 2008; Sulter, Schutte, & Miller, 1995). Although there is no evidence to suggest that a shortened method involving collection of data on semitones C, E, G, and A in each octave of a participant’s range provides data comparable to a complete VRP, many studies have adopted this method and collect fewer data points. Unfortunately, no underlying theory exists that suggests that this is the minimum number of data points allowable to adequately capture a VRP. Additionally, this method may miss important dips or leaps in sound pressure by skipping notes.
Additional Considerations

Most researchers have recorded VRPs on the vowel /a/ despite EUP recommendations to record on several vowels to obtain a wider range of information about the voice (Coleman et al., 1977; Gramming et al., 1986; Gramming et al., 1991; Lycke et al., 2012; Sanchez et al., 2014; Sulter et al., 1995). Current studies have involved researcher or clinician feedback and coaching throughout the elicitation process (Coleman et al., 1977; Gramming et al., 1991; Hallin et al., 2012; Sanchez et al., 2014; Sulter et al., 1995). Researchers and clinicians modeled different registers, different techniques, or used imagery to elicit desired semitones and intensities from participants.

Analyzing VRP Results

General Analysis

A finished VRP provides a graphic representation of voice function in terms of frequency versus intensity; however, one can also extract various numeric and mathematic parameters from the data collected. According to expert judges in one study, the shape should consistently resemble the “sum of two overlapping ellipses” (Sulter et al., 1995, p. 1078). Measures of interest in many studies include frequency and intensity minimums and maximums as well as frequency and intensity ranges (DeJonckere et al., 2003; Hallin et al., 2012; Lamarche, Ternström, & Pabon, 2010; Sanchez et al., 2014; Sulter et al., 1995).

Area

Area under the curve and above the curve appears to be the most obvious method to analyze VRP area, however, this number may not provide much meaningful information about voice function. Some studies include enclosed area, which can be calculated in different
manners, as an important measure (DeJonckere et al., 2003; Hallin et al., 2012; Lamarche et al., 2008; Lycke et al., 2012; Sanchez et al., 2014; Sulter et al., 1995; Sulter et al., 1994).

Sometimes enclosed area is divided by a constant reference range that represents normal speaking range from 40 dB to 110 dB (Sulter et al., 1994), which creates a ratio-type measure. This calculation reveals how much of the normal speaking range an individual can use based on VRP results. Because controlling loudness is especially important for singers, other studies focusing on singers have examined how much of the VRP area falls above 90 dB (Lamarche et al., 2010).

Researchers have also used enclosed area to distinguish between head and chest voice portions of the VRP (Lycke et al., 2012). Based on perceptual voice observations or a register dip present in the VRP, areas for head voice and chest voice are calculated separately. Some studies also considered register dip location to be an important parameter (Lycke et al., 2012; Sulter et al., 1995) because the upper VRP contour signals a transition from the modal register (chest voice) to the falsetto register (head voice). Some studies have indicated that this register dip may appear less pronounced in trained singers and more pronounced in untrained participants (DeJonckere et al., 2003).

**Slope**

Another measure, VRP slope, is used to analyze VRP data (Lamarche et al., 2008; Lycke et al., 2012). Researchers have applied regressions to each curve or to the overall VRP in order to determine the slope. Slopes can then be compared among participants and within and between groups. Fourier descriptors also provide important information for VRP analysis (Pabon et al., 2011; Sulter et al., 1994). Traditionally, Fourier descriptors reveal specific shape
changes by calculating the slope and angle between each data point (Pabon et al., 2011; Sulter et al., 1994). These values are then transformed to amplitude values, which reveal how much each point contributes to the VRP relative to the other points (Sulter et al., 1994).

Pabon, Ternström, and Lamarche (2011) recently proposed a new method to use Fourier descriptors for VRP analysis. This method uses the Fourier transform to up-sample and down-sample VRP contours. When additional harmonics are added with zero amplitude, the contour integrity is maintained, and additional data points are added to the VRP. This process preserves the details of the contours, smooths the contours, and allows VRPs with different numbers of data points to be more easily compared. A similar process can also be applied to down-sample the contours and remove some data points. Down-sampled contours will not have the additional details that contours with more data points contain, but down-sampling may facilitate comparison with VRPs containing fewer data points.

Clinical Utility

Together, the analysis methods and parameters discussed above can aid voice professionals in interpreting VRP results. Researchers analyze VRP results to learn more about voice function, voice gender differences, voice changes during puberty, differences between healthy and disordered voices, and differences between professional voice users and non-professional voice users (Hallin et al., 2012). Researchers have used information from VRPs to compare groups, and many have attempted to create normative data for specific groups of voice users (Hallin et al., 2012; Pabon et al., 2011; Sanchez et al., 2014; Sulter et al., 1995). One study examined differences between male and female subjects (Sanchez et al., 2014). The study
examined enclosed area, minimum and maximum-intensity and frequency, intensity range, and frequency range. The results indicated significant differences between males and females for minimum and maximum frequencies, maximum-intensity, and enclosed area.

Another study examined differences between trained and untrained voice users to establish norms (Sulter et al., 1995). Both males and females participated in the study, and the researchers examined frequency and intensity ranges, register dips, and enclosed area to find differences between groups. Results of the study indicated that trained males and females had significantly lower minimum intensities and less prominent register dips than their untrained counterparts. Furthermore, trained females had significantly larger frequency ranges and enclosed areas than untrained females. Trained males had a greater low-frequency range while untrained males had a greater high-frequency range. Although this study attempted to establish norms, researchers have commented that despite attempts to standardize VRP elicitation and analysis, the wide variety of methods employed across studies make establishing norms extremely difficult (Pabon et al., 2011).

Despite the many discrepancies in VRP elicitation and the deficiencies in strong normative data, the VRP still maintains its utility in clinic settings. Coleman (1993) states that eliciting a VRP may prove particularly helpful in clinical settings because the VRP measures how well the vocal mechanism functions as a whole. Other voice measures may capture only one aspect of voice function, such as respiratory function or laryngeal function. Clinicians have used computerized VRPs for visual biofeedback to help clients understand changes in frequency and intensity during voice therapy (Hallin et al., 2012). In addition, significant changes in VRP parameters such as minimum-intensity, enclosed area, minimum frequency, and frequency range
can determine whether clients are making progress across therapy sessions (DeJonckere et al., 2003). Some researchers even suggest that when a computerized glissando elicitation method is used, unbroken contours may indicate vocal health (Hallin et al., 2012).

The VRP also provides important information for singers. Not only can the VRP help to track changes or improvements due to vocal training, but it can also aid in voice classification, evaluation of repertoire, and identification of problematic areas in a singer’s range (Coleman, 1993; Hallin et al., 2012; Lamarche et al., 2008; Lycke et al., 2012; Lycke, Ivanova, Van Hulle, Decoster, & De Jong, 2013; Titze & Hunter, 2011). Traditionally, there are three voice classifications for males and three for females (Lycke et al., 2013). Males are classified as tenors, baritones, or basses. Females are classified as sopranos, mezzo-sopranos, or contraltos. Researchers, voice teachers, and singers agree that correct voice classification is important for optimum vocal performance and vocal health (Lycke et al., 2013). Ideally, a vocalist’s classification will allow him or her to perform primarily in the middle of his or her VRP (Titze, 2000, p. 264).

The middle region of the VRP should allow for stable phonation and equal frequency and intensity changes. However, no clear protocol exists for voice classification, and researchers question whether voices naturally fall into three categories (Lycke et al., 2013). Recent studies have evaluated the VRP as a tool for voice classification (Lycke et al., 2012; Lycke et al., 2013). These studies examined numerous VRP parameters to determine which parameter or combination of parameters would best classify voices into groups. These parameters included enclosed area, register dips, area and perimeter of head and chest voice regions, slopes of upper and lower contours, and various ratios comparing different VRP parameters (Lycke et al., 2012;
Lycke et al., 2013). For both males and females, voices were best classified in three clusters. Eighty-six percent of males were clearly placed into one of three groups based on the frequency of the register dip present in their VRPs (Lycke et al., 2013). Over 80% of females were clearly placed into 1 of 3 groups based on the ratio of the perimeter of the chest voice region versus the total VRP perimeter (Lycke et al., 2012). When singing teachers were first given the VRP results and classification for a singer, singers were more consistently classified across different voice teachers (Lycke et al., 2012). These results indicate that three different voice types may naturally exist. Although voice classification includes other factors such as timbre and physical attributes, VRP results can play an important role in making voice classification a more consistent process.

Considerations for the Current Study

Environment and Equipment

Sound level meter setting was considered as an important factor, and the EUP recommends a dB A setting to reduce environmental noise (Schutte & Seidner, 1983). To correct for background noise, the signal must be at least 10 dB greater than the ambient room noise (Šrámková, Granqvist, Herbst & Švec, 2015). Preliminary recordings of room noise were considered in order to ensure that the quietest signal was 10 dB greater than the room noise. To avoid near field distortion effects, the EUP recommends a 30 cm mouth to microphone distance. Many protocols reported in the literature used this distance in order to facilitate comparisons with other VRPs. In addition, duration of each elicitation session was considered as a way to evaluate average time differences between the full and the shortened method (see Appendix A).
**Elicitation Considerations**

Warm-ups were considered when creating a VRP protocol as warm-ups are an influential factor for VRPs (Coleman, 1993). Various vocalises\(^1\) were examined in order to determine how to best elicit the largest semitone range by semitone steps. To elicit the largest semitone range, a vocalise was sought that started at a higher semitone to stretch the vocal folds so that the “cover” would be engaged more than the “body” of the vocal folds (Titze, 2000, p. 229). Therefore, more semitones could be approximated more easily than if the exercise ascended from a lower semitone with more body-engaged vocal fold vibration. Semitone duration was also considered as an important factor in VRP elicitation (Coleman, 1993). Different voice qualities to accept were also considered as an important factor that might influence VRP shape.

The number of productions permitted or required was also considered as an influential part of the procedure as well as the amount and type of cues, modeling, coaching, and feedback (Coleman, 1993). Because this study sought to establish a broad and flexible elicitation method, all types of coaching or cueing were considered. Coaching or feedback present in the literature included the following methods: a.) Verbal encouragement that included comments such as ‘great job, try to be that quiet again,’ ‘good, try to be even quieter,’ etc. b.) Verbal instruction that included coaching the participant to use a different register (use your ‘head voice’ or ‘chest voice’) or instructing the participant to maintain the vowel, change mouth opening, move from a whisper to barely phonating, or move from *piano* to a quieter production. c.) Imagery that included verbal or nonverbal cues such as, ‘Think about going up and over the pitch,’ or using

---

\(^1\) A vocalise is “a musical passage sung upon one vowel as an exercise to develop flexibility and control of pitch and tone.” ([https://www.collinsdictionary.com/us/dictionary/english/vocalise](https://www.collinsdictionary.com/us/dictionary/english/vocalise)).
hand gestures to indicate descending or ascending loudness. Repeated productions at each sampling point were presented in the literature as an influential factor (Coleman, 1993).

For the new short method, the procedure in a pilot study by van Mersbergen and Melton (2007) was considered. They found on a full VRP protocol that the average decibel interval between each sampling point was three decibels. Therefore, they concluded that a greater than 7 decibel difference between points on this shortened method would be two times greater than the average difference expected. Therefore, additional sampling to maintain important information in the VRP contours would be required. This VRP was created starting at the participant’s lowest semitone produced and used every perfect fifth and octave throughout the participant’s semitone range (e.g. minimum of C₃ means sampling at C₃, G₃, C₄, G₄, C₅, etc.). Whenever a 7 dB or greater disparity was found between these intervals, a middle semitone was collected.

Other methods from the literature were also considered, such as the C, E, G, A method (CEGA; Gramming, Sundberg, & Akerland, 1991; Lamarche et al., 2008; Sulter, Schutte, & Miller, 1995). The CEGA VRP used the semitones, C, E, G, and A in every octave in each participant’s range. If the participant’s minimum and maximum semitones were semitones other than C, E, G, or A, these notes were also included.

**Purpose**

A Voice Range Profile provides useful information regarding vocal function, voice range, and intensity ranges for discrete semitones. This information is particularly useful to researchers, clinicians, and voice teachers. However, current methods remain largely unstandardized, and current recommendations involve specific equipment and environments to which many researchers, clinicians, and voice teachers may not have access (Pabon et al., 2011).
Current and past research studies demonstrate poor reporting of elicitation procedures and equipment used (DeJonkere et al., 2003; Gramming et al., 1986; Lycke et al., 2012). In addition, despite recommendations for standardization, environment and mouth to microphone distance vary across studies. Without standardized methods for elicitation, comparisons (particularly visual comparisons) are difficult, and one must sample intensity range at every semitone in a participant’s range to maintain VRP validity, which requires a significant time commitment that may not be practical for most users. Although current studies and common procedures suggest eliciting a VRP in an environment with low ambient noise and using high-quality sound level meters, no studies have examined whether VRP reliability can be maintained in an environment with more ambient noise if adjustments are made to correct for these issues (Coleman et al., 1977; Gramming et al., 1991; Hallin et al., 2012; Lamarche et al., 2008; Lycke et al., 2012; Sanchez et al., 2014; Schutte & Seidner, 1983; Sulter et al., 1995). Therefore, examination of VRP differences when elicited in an environment with ambient noise comparable to that of a music room or studio is important when creating a more practical VRP for everyday use in research settings, clinics, and voice studios. Furthermore, a method that involves collection of fewer data points while maintaining VRP shape and validity will make VRP elicitation more practical for frequent use in voice studios and clinics in which time constraints may be an issue.

The purpose of this study was to generate a practical protocol for Voice Range Profile (VRP) use in research, clinics, and music studios. This practical protocol will allow researchers, singing teachers, and speech-language pathologists to examine voice function and development. There was one main research question: How few data points can be collected while still maintaining VRP shape and validity? If the shortened protocols provide information and detail
comparable to a full VRP, then differences between each semitone point on the full and the shortened protocols will be similar. Similarity was defined as no greater than 2 dB and within one standard deviation from the average difference across all semitones inside the minimum and maximum curves.
CHAPTER 2: METHODS

Participants

Twenty-four singers between the ages of 18 and 80 years old were recruited from the Northern Illinois University Music Department and choirs in the surrounding Chicagoland area. Inclusion criteria consisted of a history of at least two years of private voice lessons, experience singing in a collegiate setting (e.g., university ensembles, voice major or minor, university performances, etc.), and sufficient English comprehension to follow written and verbal instructions. Inclusion criteria were set to ensure adequate vocal rapport to complete the VRP elicitation, pitch matching ability, and familiarity with vocalises and coaching techniques. Exclusion criteria included a score above 28 on the VHI (Appendix B), self-reported history of or current vocal pathology (Appendix C), or obvious speaking dysphonia based on an informal auditory-perceptual evaluation. Twelve male (average age: 26 years; age range: 19-40 years) and 12 female (average age: 32 years; age range: 19-75 years) singers were included in the study. Two of the original 14 female participants were excluded from this study. One of these participants was excluded because she was unable to return for the short protocol due to scheduling conflicts. Another female participant was excluded due to a history of laryngopharyngeal reflux that she reported caused laryngeal tightness and “cricopharyngeal

---

2 Males reported average water consumption of 54.58 ounces per day. Females reported average water consumption of 52.5 ounces per day. Males reported singing for 1.96 to 3 hours per day. Females reported singing for 2.25 to 2.49 hours per day.
spasms.” Therefore, 12 female participants were included in this study. Singers were asked to report their own fach, and of the female singers, six reported themselves as sopranos, four as mezzo-sopranos, one as alto, and one reported that she is between the soprano and mezzo-soprano fachs. Of the male singers, three reported themselves as tenors, five as baritones, three as basses, and one as a baritone-bass.

**Instrumentation**

A Tenmars TM-103 sound level meter (Tenmars Electronics Co., LTD., Taipei, Taiwan) was used to capture sound pressure level. This sound level meter was chosen because it complies with ANSI Class II standards and was under $150.00, which was considered a reasonable price. A piano or electronic keyboard was used for elicitation. A data recording sheet was used to record ambient noise, participant number, intensity level for each trial, and cueing provided. A stopwatch was used to record elicitation time. An Excel file (Microsoft, Seattle, Washington, 2016) running on a Dell Optiplex 755 (IBM, Armonk, New York) was used to plot data points, interpolate data points for the sparse, short, and CEGA versions, and to analyze data.

**Procedures**

**Full Protocol**

The participants completed a brief questionnaire requesting information about demographics, singing training, and vocal health behaviors. The participants also completed the Voice-Handicap Index (Jacobson, Johnson, Grywalski, Silbergleit, Jacobson, Benninger, & Newman, 1997) and underwent an informal auditory-perceptual evaluation to rule out any
current dysphonia. The researcher then explained to the participant what a VRP is and what the participant would be doing.

**Warm-up**

Following this explanation, the researcher warmed up with the participant for 1-3 minutes using lip or tongue trills and/or pitch glides on /u/ as needed, as warm-ups are an influential factor for VRPs (Coleman, 1993).

**Establishing Semitone Range**

The researcher demonstrated the semitone range elicitation exercise using a descending, ascending, and descending major arpeggio exercise at a moderate speed (see Figure 2). This exercise was chosen because starting at a higher semitone stretches the vocal folds and the “cover” will be engaged more than the “body” of the vocal folds (Titze, 2000, p. 229). Therefore, more semitones could be approximated more easily than if the exercise ascended from a lower semitone with more body-engaged vocal fold vibration. The participant then used the exercise to move up from a comfortable semitone by semitones to obtain the highest note in his or her range. Finally, the participant used the exercise to move down from a comfortable semitone by semitones to obtain the lowest note in his or her range. All participants completed the warm-up without incident.

![Figure 2: Descending, ascending, descending arpeggio for semitone range elicitation.](image)
**Minimum-Intensity Curve Elicitation**

Next, the minimum intensities across the participant’s frequency range were obtained. Starting one perfect fifth (7 semitones) above the lowest note obtained in the semitone range, the researcher asked the participant to produce /a/ as quietly as possible on this note. Semitones were required to last for 2-3 seconds in duration (Coleman, 1993). The researcher provided imagery on each trial and demonstrated quiet phonation if necessary. On the first production, the participant produced the semitone loudly or softly without cues or models. On all following productions, the examiner provided appropriate coaching or feedback. Because this study sought to establish a broad and flexible elicitation method, all types of coaching or cueing were employed as needed. Coaching or feedback included one or more of the following methods: a.) Verbal encouragement that included comments such as ‘great job, try to be that quiet again,’ ‘good, try to be even quieter,’ etc. b.) Verbal instruction that included coaching the participant to use a different register (use your ‘head voice’ or ‘chest voice’) or instructing the participant to maintain the vowel, change mouth opening, move from a whisper to barely phonating, or move from *piano* to a quieter production. c.) Imagery that included verbal or nonverbal cues such as, ‘Think about going up and over the pitch,’ or using hand gestures to indicate descending or ascending loudness.

At least two trials per semitone were elicited with modeling, coaching, or feedback on the second production and all subsequent productions unless the participant was demonstrating significant fatigue at the extremes of his or her range. These two productions were required to be
within two decibels of one another. For example, if the participant produced a quiet production four decibels quieter than her other productions on the third attempt, additional attempts were completed to obtain a production within two decibels of that quieter production if it was to be accepted (Coleman, 1993). A maximum of six trials per semitone were allowed before moving on to the next semitone. To ensure maximum flexibility in the protocol, all vocal qualities were accepted if the semitone was recognizable and stable per the previously stated criteria. This process was repeated descending by semitones until the lowest semitone was reached. Then, the process was repeated beginning one perfect fifth above the lowest note obtained in the semitone range and ascending until the highest semitone in the participant’s range was reached.

**Maximum-Intensity Curve Elicitation**

Next, the maximum intensities across the participant’s frequency range were obtained. Starting one perfect fifth above the lowest note obtained in the semitone range, the researcher asked the participant to produce /a/ as loudly as possible on this note. The researcher provided imagery on each trial and demonstrated loud phonation if necessary. At least two trials per semitone were elicited with modeling, coaching, or feedback on the second production and all subsequent productions. A maximum of six trials per semitone were allowed before moving on to the next semitone. Again, all vocal qualities were accepted if the pitch was recognizable and stable per the previously stated criteria. This process was repeated descending by semitones until the lowest semitone was reached. Then, the process was repeated beginning one perfect fifth above the lowest note obtained in the semitone range and ascending until the highest semitone in the participant’s range was reached. Any semitones that the participant or researcher
wanted to sample again for maximum intensities were then re-sampled. Participants were asked to return later to complete a shortened Voice Range Profile. See Appendix D for a script of the elicitation method.

**Shortened Protocols**

There were three different shortened protocols. The first and second protocols used the data collected from the full protocol. The first of the shortened protocols, the *sparse* VRP (van Mersbergen & Melton, 2007), used every perfect fifth and octave throughout the participants’ semitone ranges with additional sampling whenever a 7 dB or greater disparity was found between consecutive points (e.g. minimum of C₃ means sampling at C₃, G₃, C₄, G₄, C₅, etc.). If a 7 dB discrepancy was found, an additional point was sampled halfway between the two points. If a 7 dB or more separation existed between these additional points, extra sampling was completed by splitting the difference between the two semitones.

The second shortened protocol, the *CEGA* VRP (Gramming, Sundberg, & Akerland, 1991; Lamarche et al., 2008; Sulter, Schutte, & Miller, 1995), used the semitones C, E, G, and A in every octave in each participant’s range as well as the participant’s minimum and maximum semitones (if the minimum and/or maximum were semitones other than C, E, G, or A). This protocol was generated using data from each participant’s full VRP.

The third shortened protocol, the *short* VRP proposed in this study, was completed by all participants on a subsequent day after completion of the first protocol. Participants returned and followed the same process as the full protocol without sampling all the semitones in their range. Sampling was done at every perfect fifth and octave of each octave in the participants’ semitone
ranges with additional sampling whenever a 7 dB or greater difference was found between
consecutive points. If a seven-dB difference was found, an additional point was sampled
between the two points. Additional sampling was done if a more than 7 dB difference existed
between the additional points. This additional measure served as an inter-participant reliability
measure.

Data Collection

VRP data was recorded via pencil and paper and later entered into an Excel file
(Microsoft, Seattle, Washington, 2016). Although computerized data recording could be done,
the pencil-and-paper method for recording VRP data is more basic and accessible for most
clinicians. Graphical displays assisted in visual inspection of the data.

Analysis

Male and female VRPs were analyzed separately due to histological (cellular),
morphological (muscular), and structural (vocal tract size and shape) differences known to exist
between male and female vocal tracts (Colton, Casper & Leonard, 2011, p. 68, 408-410; Titze,
2000, p. 188-191). Because of these differences, males and females have different frequency
ranges, and males are known to have larger frequency ranges than females. A larger frequency
range would require additional time to generate any frequency-range-based measure, such as the
VRPs in this study. In addition, previous studies have shown that male and female voices are
indeed distinct, statistically different groups that should therefore be analyzed separately. In
addition, minimum and maximum curves were analyzed separately due to the different
mechanisms used for soft and loud phonation (Titze, 2000, p. 263-264).
Time Comparisons (Number of Semitones and Elicitation Times)

Male and female VRPs were analyzed separately. For each version, average number of semitones collected and average time were calculated. Time for the sparse and CEGA methods was estimated based on the full VRPs. Time was estimated by dividing full VRP total time by the number of data points elicited to find average time per data point. Average time per data point was multiplied by the number of data points to estimate the time for the CEGA and sparse VRPs. In addition, two-sample equal-variance $t$ tests were used to compare elicitation time between the full and the short protocol.

Frequency Differences

Total frequency range was calculated separately for males and females. Total range refers to the specific maximum and minimum semitones that were elicited from the entire group (e.g., males or females). In addition, frequency range differences between the short and the full protocols were examined. Minimum and maximum curves were examined separately in two different manners. First, frequency range differences were examined by subtracting total number of semitones in the short protocol by the total number of semitones in the full protocol. Second, frequency range comparisons were examined at the extremes of each participant’s range for both minimums and maximums. If the number of semitones a participant produced on the short protocol differed from the minimum or maximum semitone produced on the full protocol, these differences were recorded (e.g., on the minimum curve, a participant produced two additional semitones on the low end of her range and one fewer semitone at the high end of her range during the short protocol; this would result in a total one semitone difference between these two
(protocols.) Means and standard deviations for total number of semitone differences and specific range differences (low and high) were calculated separately for males and females. Two-sample equal-variance $t$ tests were conducted to compare range in number of semitones between the full protocol and the short protocol.

**Intensity Differences**

The full VRP was elicited on the first visit. The full VRP was graphed for visual analysis, and the numeric data was entered in a separate sheet for further analysis. This data also provided information for two other protocol versions described below.

The short VRP was elicited on the second visit, and it was linearly interpolated, to estimate omitted semitones, using an Excel program to generate decibel levels for each semitone in the range. After interpolation, the short VRP was graphed, and the numeric data was retained in a separate sheet for further analysis.

In addition to the full and short VRPs elicited during the study, two VRPs were derived from the full data. The Sparsely sampled VRP, based on the full VRP data, was sampled per the short method and linearly interpolated to estimate omitted semitones. After interpolation, the sparse VRP was graphed, and the numeric data was retained in a separate sheet for further analysis. The CEGA method was sampled from the full VRP using every C, E, G, and A semitone in the participant’s range as well as minimum and maximum semitones that were not C, E, G, A. The CEGA method was also linearly interpolated to estimate omitted semitones using an Excel file. After interpolation, the CEGA VRP was graphed, and the numeric data was retained in a separate sheet for further analysis.
To compare the VRP protocols with each other, minimum intensities at each semitone for the one VRP were subtracted from the minimum intensities of the other. Maximum-intensity SPL difference scores were obtained using the same method. Semitone sampling points that were not produced on both protocols were not included in the analysis (e.g., if a participant produced an intensity of 110 dB at D₆ on the short VRP but did not produce D₆ on the full VRP, then D₆ was not included in the analysis for this participant). Points that were sampled on one protocol but not on another would have generated large SPL difference scores that would affect averages and standard deviations. SPL difference scores at each semitone point were averaged separately for males and females to determine average difference between the protocols at each semitone point.

There were two systems of analysis to determine if the VRPs were considered similar. The first considered the decibel differences. If the semitone difference was below 2 dB, it was considered to be normal according to previous studies (Coleman, 1993). The second analysis considered the variability of the difference scores. If a semitone difference fell outside one standard deviation based on the average of all the semitone differences, it was considered to be highly variable and therefore potentially dissimilar. For the minimum curves, if shortened VRP protocols yielded difference scores that were 2 dB or one standard deviation lower, it was considered favorable because such scores indicated that the protocol in question allowed for lower minimum intensities (a larger intensity range); therefore, these scores were excluded from further calculations. These scores were excluded because this research sought only to determine whether the sparse or short protocol differed from the full protocol by providing a smaller
intensity range than the full protocol. It is beyond the scope of this study to examine whether the short protocol produces significantly larger intensity ranges than the full protocol. Likewise, for the maximum curve, if the shortened VRP protocols yielded difference scores that were 2 dB or one standard deviation higher, it was considered favorable because such scores indicated that the protocol in question allowed for greater maximum intensities (a better score); therefore, these scores were excluded from further calculations.

**Criteria for Acceptance**

Once the difference scores were calculated for each semitone, any semitone greater than 2 dB or greater than 1 standard deviation was one point (semitone) in the curve that was not similar. We chose to accept that the complete VRPs were similar if fewer than 5% of the sampling points deviated from a +/-2 dB range or a +/-1 standard deviation range. Therefore, a 95% similarity level was desired for curves to be considered similar. Male minimum curves, male maximum curves, female minimum curves, and female maximum curves were analyzed separately. The following comparisons were made: full vs. short as a reliability check comparing two VRP elicitations at two different times; sparse vs. short as a reliability check comparing the modeled shortened version with the elicited shortened version; and sparse vs. CEGA to compare the two modeled versions with each other.
CHAPTER 3: RESULTS

Time Comparisons

The average number of sampling points (semitones) produced for a male full VRP was 82.75 and the average time for completion was 39.69 minutes. The average number of sampling points (semitones) produced for a female full VRP was 76.75 and the average time for completion was 38.06 minutes (Table 1).

The average number of sampling points produced for the male’s short VRP was 30.83 and the average time it took was 17.21 minutes. The average number of sampling points produced for the female’s short VRP was 28.33 and the average time for completion was 16.38 minutes.

The estimated average number of sampling points produced for the male’s sparse VRP was 29.82 and the average time it took was 14.35 minutes. The estimated average number of sampling points produced for the female’s sparse VRP was 26.67 and the average time for completion was 13.45 minutes.

The estimated average number of sampling points produced for the male’s CEGA VRP was 29.33 and the average time it took was 13.95 minutes. The estimated average number of sampling points produced for the female’s CEGA VRP was 28.55 and the average time for completion was 11.31 minutes.
For males, a two-sample equal-variance $t$ test showed that total time for the full protocol ($M=2381.17$, $SD=413.61$) was significantly different from total time for the short protocol ($M=1032.58$, $SD=185.92$), $t(22)=10.31$, $p<0.001$ (see Table 2). For females, a two-sample equal-variance $t$ test showed that total time for the full protocol ($M=2321.00$, $SD=481.14$) was significantly different from total time for the short protocol ($M=982.58$, $SD=222.37$), $t(22)=8.75$, $p<0.001$ (see Table 2).

Table 1: Male and female average elicitation time, number of sampling points (minimum plus maximum curve), and time per sampling point.

<table>
<thead>
<tr>
<th></th>
<th>Average elicitiation time, number of sampling points, and time per sampling point</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Average Minutes</td>
</tr>
<tr>
<td>Full</td>
<td>Mean 39.69</td>
</tr>
<tr>
<td></td>
<td>SD 6.6</td>
</tr>
<tr>
<td>Short</td>
<td>Mean 17.21</td>
</tr>
<tr>
<td></td>
<td>SD 3</td>
</tr>
<tr>
<td>Sparse*</td>
<td>Mean 14.35</td>
</tr>
<tr>
<td></td>
<td>SD 1.4</td>
</tr>
<tr>
<td>CEGA*</td>
<td>Mean 13.95</td>
</tr>
<tr>
<td></td>
<td>SD 1.3</td>
</tr>
</tbody>
</table>

*calculations based on Average # of semitones from the Full VRP

Table 2: Male and female two-sample equal-variance $t$ tests comparing full and short elicitation time.

<table>
<thead>
<tr>
<th></th>
<th>T-test for Time</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p$ value</td>
<td>$t$</td>
</tr>
<tr>
<td>Female Full vs. Short</td>
<td>0.00</td>
<td>8.75</td>
</tr>
<tr>
<td>Male Full vs. Short</td>
<td>0.00</td>
<td>10.31</td>
</tr>
</tbody>
</table>
Full vs. Short Comparison

Male Frequency Differences

Average semitone range differences did not exceed two semitones, although one participant produced 16 additional semitones during the short protocol (see Tables 2, 3, and 4). Average male semitone range on the minimum curve included 1.67 additional semitones during the short protocol (SD=4.52). Average male semitone range on the maximum curve included 0.5 additional semitones during the short protocol (SD=2.25). A two-sample equal-variance t test showed that frequency range on the minimum curve during the full protocol (M=39.67, SD=3.47) was not significantly different from minimum curve frequency range during the short protocol (M=41.33, SD=4.38), t(22)=1.03, p=0.31. A two-sample equal variance t test showed that frequency range on the maximum curve during the full protocol (M=43.08, SD=4.29) was not significantly different from maximum curve frequency range during the short protocol (M=43.58, SD=3.78), t(22)=0.30, p=0.77.

Table 3: Average range differences at the low and high ends of the range between the short and full protocol for male and females.
Table 4: Average frequency range differences by number of semitones between the short and the full protocol for males and females.

<table>
<thead>
<tr>
<th></th>
<th>Minimum curve</th>
<th>Maximum curve</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Mean</td>
<td>1.67</td>
<td>0.08</td>
</tr>
<tr>
<td>SD</td>
<td>4.52</td>
<td>0.76</td>
</tr>
</tbody>
</table>

**Female Frequency Differences.**

Average female semitone range on the minimum curve included 0.08 additional semitones during the short protocol ($SD=0.76$; see Tables 2, 3, and 4). Average female semitone range on the maximum curve included 0.58 fewer semitones during the short protocol ($SD=1.04$). A two-sample equal-variance $t$ test showed that frequency range on the minimum curve during the full protocol ($M=37.83$, $SD=1.90$) was not significantly different from minimum curve frequency range during the short protocol ($M=37.92$, $SD=1.98$), $t(22)=0.11$, $p=0.92$ (see Tables 2-5). A two-sample equal-variance $t$ test showed that frequency range on the maximum curve during the full protocol ($M=38.92$, $SD=1.78$) was not significantly different from maximum curve frequency range during the short protocol ($M=38.33$, $SD=1.72$), $t(22)=0.82$, $p=0.42$ (see Tables 2-5).
Table 5: Male and female two-sample equal variance \( t \) tests comparing full and short protocols on the minimum and maximum curves.

<table>
<thead>
<tr>
<th></th>
<th>p value</th>
<th>t</th>
<th>df</th>
<th>M 1</th>
<th>M 2</th>
<th>SD 1</th>
<th>SD 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female Min Full vs. Short</td>
<td>0.92</td>
<td>0.11</td>
<td>22.00</td>
<td>37.83</td>
<td>37.92</td>
<td>1.90</td>
<td>1.98</td>
</tr>
<tr>
<td>Female Max Full vs. Short</td>
<td>0.42</td>
<td>0.82</td>
<td>22.00</td>
<td>38.92</td>
<td>38.33</td>
<td>1.78</td>
<td>1.72</td>
</tr>
<tr>
<td>Male Min Full vs. Short</td>
<td>0.31</td>
<td>1.03</td>
<td>22.00</td>
<td>39.67</td>
<td>41.33</td>
<td>3.47</td>
<td>4.38</td>
</tr>
<tr>
<td>Male Max Full vs. Short</td>
<td>0.76</td>
<td>0.30</td>
<td>22.00</td>
<td>43.08</td>
<td>43.58</td>
<td>4.29</td>
<td>3.78</td>
</tr>
</tbody>
</table>

**Male Intensity Differences**

For males, the average minimum-intensity difference was 0.47 dB with a standard deviation of 6.2 dB. The average maximum-intensity difference was -0.86 dB with a standard deviation of 4.45 dB. On average, the short method captured slightly lower minimum intensities and slightly higher maximum intensities than the full method (see Figure 3). Lower minimum intensities and higher maximum intensities reflect better vocal control and are therefore desirable.
Figure 3: Average male full VRP graph compared to average male short VRP graph.

2 dB Differences

**Minimum Curve.** Sixteen out of the 46 sampling points (35%) fell outside the +/-2 dB range suggested for comparable VRPs. However, 10 of these 16 were positive, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. The remaining 13% of the sampling points were negative and outside of the 2 dB difference score range, which indicates that there was an 87% similarity in the minimum-intensity curve, which was lower than the desired 95%
similarity. The 87% similarity indicates that the short protocol and the full protocol produced different minimum curves for males.

**Maximum Curve.** Ten out of 50 sampling points (20%) were outside the +/-2 dB range suggested for comparable VRPs. However, 9 of these 10 were negative, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. This left 2% of the mean SPL difference scores that were positive and outside of the 2 dB difference score range. Therefore, there was 98% similarity between the two protocols for the maximum curve. The 98% similarity indicates that the short protocol and the full protocol produced similar maximum curves for males.

**Standard Deviation Differences**

**Minimum Curve.** Three out of 46 sampling points (7%) fell outside the one standard deviation range, which suggests that there was some variability in this comparison. However, one of these three was positive, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. This left 4% of the standard deviation scores that were negative and outside of the one standard deviation range. Therefore, there was 96% similarity between the two protocols for the minimum curve. The 96% similarity indicates that the short protocol and the full protocol produced similar minimum curves for males (see Figure 4).
Maximum Curve. Two out of 50 sampling points (4%) fell outside the one standard deviation range. Of these two, both were negative, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. This left 0% of the standard deviation scores that were positive and outside of the one standard deviation range. Therefore, there was 100% similarity between the two protocols for the maximum curve. The 100% similarity indicates that the short protocol and the full protocol produced similar maximum curves for males (see Figure 4).
Figure 4: Average male full versus short SPL difference scores across frequency range. A positive score for the minimum curve was considered favorable, and a negative score for the maximum curve was considered favorable. Favorable scores indicated that the short protocol provided lower minimum intensities and/or higher maximum intensities than the full protocol.

**Female Intensity Differences**

For females, the average minimum-intensity difference was 1.31 dB with a standard deviation of 4.66 dB and the average maximum-intensity difference was -0.25 dB with a standard deviation of 4.15 dB, which suggests that on average, the short method captured slightly lower minimum intensities and slightly higher maximum intensities than the full method (Figure 5).
Figure 5: Average female full VRP graph compared to average female short VRP graph.

2 dB Differences

**Minimum Curve.** Fourteen out of 43 (33%) sampling points were outside the +/-2 dB range suggested for comparable VRPs. All 14 were positive, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. Zero percent of the mean SPL difference scores were negative and outside of the 2 dB difference score range. Therefore, there was 100% similarity between the two protocols for the minimum curve. The 100% similarity indicates that the short protocol and the full protocol produced similar minimum curves for females.
**Maximum Curve.** Five out of 44 (11%) sampling points were outside the +/-2 dB range suggested for comparable VRPs. However, three of the five were negative, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. Therefore, 5% of the mean SPL difference scores were positive and outside of the 2 dB difference score range. Therefore, there was 95% similarity between the two protocols for the minimum curve. The 95% similarity indicates that the short protocol and the full protocol produced similar minimum curves for females.

**Standard Deviation Differences**

**Minimum Curve.** Two out of 43 (4.6%) sampling points fell outside the one standard deviation range. Of these 2, both were positive, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. Zero percent of the standard deviation scores were negative and outside of the one standard deviation range. Therefore, there was 100% similarity between the two protocols for the minimum curve. The 100% similarity indicates that the short protocol and the full protocol produced similar minimum curves for females (see Figure 6).

**Maximum Curve.** One out of 44 (2.3%) sampling points fell outside the one standard deviation range. Of this one, zero were negative, and no points were excluded from the 95% similarity calculation. Two percent of the standard deviation scores were positive and outside of
the one standard deviation range. Therefore, there was 98% similarity between the two protocols for the maximum curve. The 98% similarity indicates that the short protocol and the full protocol produced similar maximum curves for females (see Figure 6).

![Female Full Vs. Short Difference Scores](image)

Figure 6: Average female full versus short SPL difference scores across frequency range. A positive score for the minimum curve was considered favorable, and a negative score for the maximum curve was considered favorable. Favorable scores indicated that the short protocol provided lower minimum intensities and/or higher maximum intensities than the full protocol.

Points at which the short protocol provided lower minimums or higher maximums were excluded from analysis because these points were considered a favorable difference. Because of this, the Short method appears to have captured a slightly larger intensity range than the Full method while maintaining adequate similarity to the full protocol. In fact, two participants produced greater than two additional semitones at the high end of their range during the short
protocol. Overall, average SPL difference scores for male minimums ($M$: 0.47, $SD$: 6.2) and maximums ($M$: -0.86, $SD$: 4.45) as well as female minimums ($M$: 1.31, $SD$: 4.66) and maximums ($M$: -0.25, $SD$: 4.15) were less than 2 dB, which is indicative of similar scores for each sampling point between the two methods. For the +/-2 dB comparison, the short and the full protocols maintained 95% or greater similarity for male maximum curves, female minimum curves, and female maximum curves. Male minimum curves fell outside the standard for acceptability. For the +/- 1 SD comparison, the short and the full protocols maintained 95% or greater similarity for male minimum curves, male maximum curves, female minimum curves, and female maximum curves. This suggests that the full and the short protocols are similar but differ slightly for males’ minimum-intensity curves, where the full protocol yields a slightly better curve.

**Sparse vs. Short Comparison**

**Frequency Differences for Males and Females**

Because the sparse sampling was based on the full sampling, the frequency difference for the comparison sparse vs. short are the same as full vs. short. Please refer to Tables 2-4 for numerical data.

**Male Intensity Differences**

For males, the average minimum-intensity difference was 0.53 dB with a standard deviation of 5.90 dB. For males, the average maximum-intensity difference was -1.11 dB with a standard deviation of 4.16 dB. For both comparisons, the short method captured slightly captured slightly lower minimum intensities and slightly higher maximum intensities than the sparse method (see Figure 7).
Figure 7: Average male sparse VRP graph compared to average male short VRP graph.

**2 dB Differences**

**Minimum Curve.** Seventeen out of 48 (35%) sampling points were outside the +/-2 dB range suggested for comparable VRPs. However, 12 of these 17 were positive, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. Ten percent of the mean SPL difference scores were negative and outside of the 2 dB difference score range. Therefore, there was only 90% similarity between the two protocols for the minimum curve. The 90% similarity is below
the desired 95% similarity level and indicates that the short protocol and the sparse protocol produced different minimum curves for males.

**Maximum Curve.** Six out of 50 (12%) sampling points were outside the $\pm 2$ dB range suggested for comparable VRPs. All of these scores were negative, which that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. Zero percent of the mean SPL difference scores were positive and outside of the 2 dB difference score range. Therefore, there was 100% similarity between the two protocols for the maximum curve. The 100% similarity indicates that the short protocol and the sparse protocol produced similar maximum curves for males.

**Standard Deviation Differences**

**Minimum Curve.** Three out of 48 (6%) sampling points fell outside the one standard deviation range. One of the three was positive, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. Four percent of the standard deviation scores were negative and outside of the one standard deviation range. Therefore, there was 96% similarity between the two protocols for the minimum curve. The 96% similarity indicates that the short protocol and the sparse protocol produced similar minimum curves for males (see Figure 8).

**Maximum Curve.** One out of 50 (2%) sampling points fell outside the one standard deviation range. This point was negative, which indicated that the short protocol captured more
of an individual’s intensity range, and this point was excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. Zero percent of the standard deviation scores were positive and outside of the one standard deviation range. Therefore, there was 100% similarity between the two protocols for the maximum curve. The 100% similarity indicates that the short protocol and the sparse protocol produced similar maximum curves for males (see Figure 8).

Figure 8: Average male sparse versus short SPL difference scores across frequency range. A positive score for the minimum curve was considered favorable, and a negative score for the maximum curve was considered favorable. Favorable scores indicated that the short protocol provided lower minimum intensities and/or higher maximum intensities than the sparse protocol.
Female Intensity Differences

For females, the average minimum-intensity difference was 1.58 dB with a standard deviation of 4.21 dB, and the average maximum-intensity difference was -0.37 dB with a standard deviation of 4.09 dB. These comparisons suggest that the short method captured slightly lower minimum intensities and slightly higher maximum intensities than the sparse method (see Figure 9).

Figure 9: Average female sparse VRP graph compared to average female short VRP graph.

2 dB Differences

Minimum Curve. Thirteen out of 43 (30%) sampling points were outside the +/-2 dB range suggested for comparable VRPs. All of these points were positive, which indicated that
the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. Zero percent of the mean SPL difference scores were negative and outside of the 2 dB difference score range. Therefore, there was 100% similarity between the two protocols for the minimum curve. The 100% similarity indicates that the short protocol and the sparse protocol produced similar minimum curves for females.

**Maximum Curve.** Nine out of 44 (20%) sampling points were outside the +/-2 dB range suggested for comparable VRPs. However, eight of the nine points were negative, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s intensity range. Two percent of the mean SPL difference scores were positive and outside of the 2 dB difference score range. Therefore, there was 98% similarity between the two protocols for the maximum curve. The 98% similarity indicates that the short protocol and the sparse protocol produced similar maximum curves for females.

**Standard Deviation Differences**

**Minimum Curve.** Two out of 43 (5%) semitones fell outside the one standard deviation range. Both were positive, which indicated that the short protocol captured more of an individual’s intensity range, and these points were excluded from the final calculation because this study focused on instances in which the short protocol captured less of an individual’s
intensity range. Zero percent (0%) of the standard deviation scores were negative and outside of the one standard deviation range. Therefore, there was 100% similarity between the two protocols for the minimum curve. The 100% similarity indicates that the short protocol and the sparse protocol produced similar minimum curves for females (see Figure 10).

**Maximum Curve.** One out of 44 semitones (2%) fell outside the one standard deviation range. None were negative, so, no points were excluded from the 95% similarity calculation. Two percent (2%) of the standard deviation scores were positive and outside of the one standard deviation range. Therefore, there was 98% similarity between the two protocols for the maximum curve. The 98% similarity indicates that the short protocol and the sparse protocol produced similar maximum curves for females (see Figure 10).
Figure 10: Average female sparse versus short SPL difference scores across frequency range. A positive score for the minimum curve was considered favorable, and a negative score for the maximum curve was considered favorable. Favorable scores indicated that the short protocol provided lower minimum intensities and/or higher maximum intensities than the sparse protocol.

Results revealed that the sparse and short methods generated very similar VRPs. Average SPL difference scores for male minimums ($M: 0.53, SD: 5.90$) and maximums ($M: -1.11, SD: 4.16$) as well as female minimums ($M: 1.58, SD: 4.21$) and maximums ($M: -0.37, SD: 4.09$) were less than 2 dB. For the +/-2 dB comparison, the short and the full protocols maintained 95% or greater similarity for male maximum curves, female minimum curves, and female maximum curves. Male minimum curves fell outside the standard for acceptability. Furthermore, the +/- 1 SD comparison showed that the short and the full protocols maintained 95% or greater similarity for male minimum curves, male maximum curves, female minimum curves, and female
maximum curves. These results were expected because both protocols used the same method. Additionally, these results mirror the previous full vs short comparison because the data for the full and the sparse comparisons are essentially the same.

**Sparse vs. CEGA Comparison**

**Frequency Differences for Males and Females**

Because the sparse and CEGA sampling were based on the full sampling, the frequency ranges are the same and therefore there are no frequency differences for this comparison.

**Male Intensity Differences**

For males, the average minimum-intensity difference was 0.07 dB with a standard deviation of 2.03 dB, which suggests that the CEGA method captured slightly more of an individual’s VRP range. For males, the average maximum-intensity difference was -0.06 dB with a standard deviation of 2.22 dB, which suggests that the CEGA method captured slightly lower minimum intensities and slightly higher maximum intensities than the sparse method (see Figure 11).
Figure 11: Average male sparse VRP graph compared to average male CEGA VRP graph.

2 dB Differences

Minimum Curve. Two out of 48 (4%) sampling points were outside the +/-2 dB range suggested for comparable VRPs. One was negative, which indicated that the sparse VRP provided more information, and this point was therefore excluded from the 95% similarity calculation. Two percent of the mean SPL difference scores were positive and outside of the 2 dB difference score range. Therefore, there was 98% similarity between the two protocols for the minimum curve. The 98% similarity indicates that the CEGA protocol and the sparse protocol produced similar minimum curves for males.
**Maximum Curve.** Three out of 52 (6%) sampling points were outside the +/-2 dB range suggested for comparable VRPs. One of the three was positive, which indicated that the sparse VRP provided more information, and this point was therefore excluded from the 95% similarity calculation. Three percent of the mean SPL difference scores were negative and outside of the 2 dB difference score range. Therefore, there was 97% similarity between the two protocols for the maximum curve. The 97% similarity indicates that the CEGA protocol and the sparse protocol produced similar maximum curves for males.

**Standard Deviation Differences**

**Minimum Curve.** Two out of 48 (4%) semitones fell outside the one standard deviation range. One was negative, which indicated that the sparse protocol provided more information, and this point was therefore excluded from the 95% similarity calculation. Two percent of the standard deviation scores were positive and outside of the one standard deviation range. Therefore, there was 98% similarity between the two protocols for the minimum curve. The 98% similarity indicates that the CEGA protocol and the sparse protocol produced similar minimum curves for males (see Figure 12).

**Maximum Curve.** Two out of 52 (4%) semitones fell outside the one standard deviation range. One was positive, which indicated that the sparse protocol provided more information, and this point was therefore excluded from the 95% similarity calculation. Two percent (2%) of the standard deviation scores were negative and outside of the one standard deviation range. Therefore, there was 98% similarity between the two protocols for the maximum curve. The 98%
similarity indicates that the CEGA protocol and the sparse protocol produced similar maximum curves for males (see Figure 12).

![Male Sparse vs CEGA Difference Scores](image)

**Figure 12:** Average male sparse versus CEGA SPL difference scores across frequency range. A negative score for the minimum curve was considered favorable, and a positive score for the maximum curve was considered favorable. Favorable scores indicated that the sparse protocol provided lower minimum intensities and/or higher maximum intensities than the CEGA protocol.

**Female Intensity Differences**

For females, the average minimum-intensity difference was 0.04 dB with a standard deviation of 2.89 dB, which suggests that the CEGA method captured slightly more of an individual’s VRP range. For females, the average maximum-intensity difference was -0.01 dB with a standard deviation of 2.38 dB, which suggests that the CEGA method captured slightly
lower minimum intensities and slightly higher maximum intensities than the sparse method (see Figure 13).

![Female Sparse vs. CEGA Average VRP](image)

**Figure 13:** Average female sparse VRP graph compared to average female CEGA VRP graph.

### 2 dB Differences

**Minimum Curve.** No sampling points (0%) were outside the +/-2 dB range suggested for comparable VRPs. Therefore, no points were excluded from the 95% similarity calculation. Zero percent of the mean SPL difference scores were positive and outside of the 2 dB difference score range. Therefore, there was 100% similarity between the two protocols for the minimum curve. The 100% similarity indicates that the CEGA protocol and the sparse protocol produced similar minimum curves for females.
Maximum Curve. One out of 45 (2%) sampling points were outside the +/-2 dB range suggested for comparable VRPs. Of these, none were positive, so no points were excluded from the 95% similarity calculation. Two percent of the mean SPL difference scores were negative and outside of the 2 dB difference score range. Therefore, there was 98% similarity between the two protocols for the maximum curve. The 98% similarity indicates that the CEGA protocol and the sparse protocol produced similar maximum curves for females.

Standard Deviation Differences

Minimum Curve. No semitones (0%) fell outside the one standard deviation range. Therefore, no points were excluded from the 95% similarity calculation. Zero percent (0%) of the standard deviation scores were positive and outside of the one standard deviation range. Therefore, there was 100% similarity between the two protocols for the minimum curve. The 100% similarity indicates that the CEGA protocol and the sparse protocol produced similar minimum curves for females (see Figure 14).

Maximum Curve. No semitones (0%) fell outside the one standard deviation range. Therefore, no points were excluded from the 95% similarity calculation. Zero percent of the standard deviation scores were negative and outside of the one standard deviation range. Therefore, there was 100% similarity between the two protocols for the maximum curve. The 100% similarity indicates that the CEGA protocol and the sparse protocol produced similar maximum curves for females (see Figure 14).
Figure 14: Average female sparse versus CEGA SPL difference scores across frequency range. A negative score for the minimum curve was considered favorable, and a positive score for the maximum curve was considered favorable. Favorable scores indicated that the sparse protocol provided lower minimum intensities and/or higher maximum intensities than the CEGA protocol.

Results indicate that the sparse and CEGA methods generated very similar VRPs. Average SPL difference scores for male minimums ($M: 0.07$, $SD: 2.03$) and maximums ($M: -0.06$, $SD: 2.22$) as well as female minimums ($M: 0.04$, $SD: 2.89$) and maximums ($M: -0.01$, $SD: 2.38$) were less than 2 dB, which is indicative of similar scores for each sampling point between the two methods. The sparse and the CEGA protocols maintained 95% or greater similarity for male minimum curves, male maximum curves, female minimum curves, and female maximum curves for both the decibel comparison and the standard deviation comparison. These results were expected because both the sparse method and the CEGA method were modeled using the
full data. However, it appears that the sparse sampling produces a larger VRP for males and the CEGA sampling produces a larger VRP for females.
CHAPTER 4: DISCUSSION

This study sought to evaluate whether a shortened protocol for VRPs would reflect similar voice functioning to allow researchers, speech-language pathologists, and singing teachers to more easily examine vocal function and development. To do this, two shortened VRPs were modeled using data from participants’ full VRPs elicited in realistic environments (e.g., music studio, office, church, clinic room). The sparse VRP was based on the method previously proposed by van Mersbergen and Melton (2007), and the CEGA VRP was based on the method used by Sulter, Schutte, and Miller (1995). Finally, the short protocol, based on van Mersbergen and Melton’s protocol (2007), was elicited from each participant. The main research question examined how few data points could be collected while maintaining VRP shape and validity.

Both the sparse VRP and the CEGA VRP collected from the full VRP required fewer data points than the full VRP and did not significantly differ from each other. Per mean SPL difference scores, both shortened methods appeared to generate results comparable to those generated with the Full method. In addition, both shortened methods were estimated to require fewer minutes to generate than did the Full method. When participants returned a second time to produce the short VRP, several of the participants produced larger frequency ranges compared with the full VRP elicitation, and this effect was most evident in the highest semitones. Because the highest semitones at maximum-intensity were elicited last, this finding may be attributed to a
fatigue effect during the full VRP. These results indicate that the shortened methods maintain the validity of a full VRP and may have several advantages compared to a traditional full VRP, such as fewer data points, shorter elicitation time, and larger frequency ranges.

**Elicitation Considerations**

While the short VRP appeared to be more efficient than the full VRP, it also demonstrated some shortcomings. When there was a 7 dB difference between neighboring frequencies, multi-semitone jumps and additional sampling appeared to generate additional difficulty for participants around register shifts or at the extreme high end of a participant’s range. Based on informal observations, when a participant was required to make a register shift from one frequency sampling point to the next, the back and forth shifting between registers presented difficulties to the participant and often required additional trials and maximal cueing on these semitones. The additional sampling presumably disrupted the predictable ascending pattern that the participant had been expecting. When additional sampling was not required near register shifts, these areas appeared to present less difficulty, requiring fewer trials and less cueing than they would if additional sampling were required. This difficulty was not present during the full VRP, and participants appeared to find it easier to move in predictable single-semitone increments near register shifts because of the reduced technical load. Although this characteristic may present unnecessary difficulty for non-singers, it might prove useful for assessing vocal precision and control in singers. However, this would also demand that additional data be included in the VRP, such as spectral data (Gramming et al., 1986) or register data (Lamesch et al., 2012).
Informal observations revealed that participants appeared to need about the same amount of cueing for the minimum and maximum curves. Cueing was provided most on the first few semitones, at the “lowest” end of the range, and at the “highest” end of the range. One area of difference was on the minimum curve at the “high” end of the range. Many participants needed extra cueing and reminders to produce these “high” notes as softly as they could, even if the intensity level was not as low as previous lower notes. This was especially evident during the shortened protocol due to multiple-semitone jumps. Overall, the amount of cueing needed depended on the participant’s comfort level and vocal control—the better the participant’s comfort level and vocal control, the fewer cues needed.

Another shortcoming of the short protocol was the demand placed on the elicitor. Because the protocol was comprehensive, increased training in the procedure would be recommended. Demands of the full protocol included data gathering, providing cues, imagery, verbal encouragement, and providing appropriate semitones on the keyboard for each trial. Basic keyboard proficiency would be required. Additionally, the shortened protocol included the additional task of noticing 7 dB differences and choosing the appropriate sampling point between these differences. A knowledge of basic music theory is also necessary to select the set sampling points based on the participant’s frequency range (e.g., octave and perfect fifth from the base octave). Not all clinicians and researchers will have adequate piano skills and music theory knowledge for this method. Despite these shortcomings, most singing teachers and voice speech-language pathologists with music backgrounds will already possess this knowledge and be comfortable using these tools. Even so, it appears as though a shortened version does capture
frequency and intensity range more completely because its shortened time mitigates vocal fatigue and mental concentration.

Because the short method had observed drawbacks, the CEGA method might be a more efficient method to capture a shortened VRP. Unfortunately, this study did not directly test the short version and the CEGA version. Rather, the sparse version and the CEGA version were compared because both were based on data collected from the full VRP revealing negligible frequency and intensity differences and direct comparisons between the short and the CEGA versions were not appropriate. Nevertheless, the short/sparse and CEGA versions can be compared conceptually. There are a few differences between the short/sparse method and the CEGA method. The first is that the short/sparse method is participant driven. This means that data points on the VRP are based on the participant’s frequency range and vocal capabilities. The CEGA, on the other hand, had distinct data points that all participants produced. With the same semitones sampled for each participant, the CEGA protocol provides structure that can facilitate comparison across participants, which would be useful in research seeking to compare VRPs across participants. However, a participant-driven method with additional sampling as needed could reveal important areas of intensity change in a participant’s range (e.g., register shifts, problematic areas for singers, etc.) that might be useful for determining a singer’s fach or a participant’s vocal health.

The perfect fifth and octave intervals were chosen for the short/sparse method because a perfect fifth is a common interval that should facilitate pitch matching, which has been known to affect VRP shape (Coleman, 1993). Pitch matching ability is known to influence vocal technique
because if a person is having difficulty matching pitch, then it can be presumed that more
cognitive resources are being allocated to matching the pitch than to vocal technique (Alain,
2007). If vocal technique is poor, a person may not produce a VRP that is most representative of
his or her current voice function and abilities. Pitch matching ability is better when the individual
is more familiar with the chord structure being used (Alain, 2007; McLachlan, Marco & Wilson,
2013). Therefore, common chords will facilitate pitch matching abilities for those with little or
no music training (McLachlan, Marco, Light & Wilson, 2013; McLachlan, Marco & Wilson,
2013). The Sparse/short protocol sought to facilitate pitch matching by using two common
intervals, the perfect fifth and the octave. Unlike the Sparse/short protocol, the CEGA method
may be difficult for non-singers and those with pitch-matching difficulty because it is based on a
major chord with an additional sixth tone.

Because results of this study appear to indicate that rigid sampling is beneficial for both
participant comfort and ease of elicitation, a method similar to the CEGA method that uses a
more familiar chord type, such as a major seventh, might be easier for non-singers and those with
pitch matching difficulty in order to combine the advantages of rigid sampling with fewer data
points. In addition, because chord familiarity is known to facilitate pitch matching, future
protocols might increase chord familiarity by playing the entire chord every four semitones to
increase participant familiarity of the chord and ideally help the participant in motor planning
(Alain, 2007; McLachlan, Marco & Wilson, 2013).
Limitations to the Research

One limitation present in this research study was the number of participants. Twelve male and 12 female participants completed the study. A greater number of participants might show greater difference between the two estimated shortened protocols and therefore could address some of the aforementioned questions. In addition, most participants were between 20 and 30 years of age. A larger sample size might have allowed for a more representative age range. Furthermore, larger numbers might have uncovered fach differences within each sex. The small sample size also influenced difference scores at extremes of the range. For example, only two participants produced $F_6$, so standard deviation scores at this point were larger than at semitone points that all participants produced. Estimations on how many more participants would be required for this additional information are unknown and cannot be gleaned from the low numbers in this study. However, other studies in the literature included 20-30 participants (DeJonckere et al., 2003; Gramming et al., 1991; Hallin et al., 2012; Lamarche et al., 2008; Lamesch et al., 2012), and a few studies used 200-260 participants (Lycke et al., 2012; Lycke et al., 2013; Sulter et al., 1995).

An additional limitation was participant self-report of fach. Several participants reported that they were sopranos, yet these participants had difficulty producing semitones above $B_5$ or $C_6$. These participants were either choosing to sing outside of their fachs, were misinformed about their own fachs, or did not have sufficient training for a fach to be properly assigned. Although self-report of fach made it difficult to obtain a sample equally representing all the
major fachs (e.g., soprano, mezzo-soprano, alto, tenor, baritone, bass), the VRP may be a useful tool in helping to place singers in an appropriate fach.

Another limitation was that the CEGA VRP was modeled from the full VRP and not actually elicited. Although we can compare the estimated versions, an elicited CEGA VRP would allow for more accurate comparison of total elicitation time, ease of elicitation, and participant comfort (e.g., compared with difficult jumps at register shifts noted on the short protocol described above) which cannot be assessed with modeling alone. Therefore, we cannot determine, based on the results of this study, whether the CEGA method or the short method would be more practical and beneficial.

An added limitation was higher than expected ambient noise present in one of the elicitation rooms. Ambient noise was acceptable if it was below 40 dB A weighted. One elicitation room’s ambient noise remained between 40 and 43 dB A. Although higher ambient noise is undesirable, an environment with higher ambient noise is a realistic environment that will be encountered in many settings, including clinical practice. The louder ambient noise may or may not have affected the level meter readings, but most likely the change would be negligible from a recording standpoint. It is more likely that a louder environment would cause participants to produce louder phonations due to the Lombard effect (Zollinger & Brumm, 2011). This would ultimately affect the minimum-intensity curve. Furthermore, ambient noise may affect the participant’s ability to hear production of soft signals and therefore influence the soft phonation contours of the VRP. More information is needed to determine how much ambient noise influences VRPs.
A final limitation in this study was the researcher’s comfort level while eliciting the protocol. Due to initial unfamiliarity with VRP elicitation as well as the additional steps required during the short protocol, initial elicitation was difficult. This may have increased total time for the VRPs with the first few participants. However, after completing several participants’ VRPs, the researcher’s comfort and elicitation speed increased. Therefore, later VRP elicitation times were likely not affected by level of researcher familiarity and comfort with elicitation. Clinician comfort level is likely to be influenced by personal factors, such as keyboard skills, knowledge of music theory, clinical experience, and knowledge of vocal coaching and technique.³

**Future Directions**

To facilitate more reliable comparison in the future, the CEGA protocol should be elicited and compared to the short method described and elicited here. Perhaps these should be compared with a new rigid sampling method that considers the strengths and weaknesses present in the CEGA and Short methods. A new rigid sampling method might decrease cognitive load on the clinician (rigid sampling eliminates a need to attend to 7 dB differences and select additional sampling points) and increase participant comfort (rigid sampling presents a predictable pattern), select musical intervals that will facilitate easy pitch matching for non-singers, and maintain about the same number of semitone sampling points used in the CEGA and short methods.

---

³ The researcher in this study had minimal keyboarding skills and very basic music theory knowledge, and she became comfortable after eliciting 10-15 full and 10-15 short VRPs.
In addition, future research should use a larger sample size. A larger sample size should include singers of different fachs and non-singers to determine how useful a shorter VRP would be with patients who are non-singers as well as those who are singers. Of course, this might require multiple studies addressing these questions individually.

For singers, whose complaints are often more specific and less noticeable than non-singers, additional measures could be added to a VRP to more closely examine vocal precision and control. These measures might include duration for which a semitone can be sustained (Titze, 2016), auditory-perceptual ratings (Lamesch et al., 2012), number of trials needed per sampling point, number of cues needed, participant perception of task difficulty and comfort during the task, etc. Although the shortened elicitation method appeared to be more difficult for participants due to multi-semitone jumps and additional sampling, it might have potential to provide additional information regarding a participant’s vocal precision, control, and technique that could prove valuable when evaluating singers.

Future studies might examine whether the shortened method can provide this information and how to measure these differences (e.g., number of trials needed, clinician rating scales, participant effort or comfort scales, number of cues needed, etc.). In order to determine if new protocols are indeed practical for everyday use, additional research might examine participant comfort and effort through a survey taken after elicitation. In addition, the clinician’s cognitive load could be examined by asking clinicians or voice teachers who are unfamiliar with the procedure to elicit several VRPs and rate their comfort, cognitive load, and confidence after eliciting the VRP using the specified method. Furthermore, additional studies might examine the
amount of instruction required for clinicians, voice teachers, and researchers to be reliable in this procedure.

Future studies should also include clinicians who are non-singers and/or non-musicians to determine whether the protocol is easily understandable and usable for those who do not have an extensive background in music or music theory. Because the VRP in general as well as the short protocol proposed in this study require specialized knowledge for elicitation and interpretation, future studies should also examine training methods for both elicitation and interpretation of results. Formalized training will be necessary not only for proper interpretation of results but also for maintaining standardization of the process. Standardization will allow for reliable comparisons across VRPs.

Finally, the effects of ambient noise that might be present in realistic elicitation settings on the VRP and the participant should be examined. Ambient noise could be examined by comparing several VRPs elicited in varying levels of ambient noise within participants to determine whether the ambient noise affects the VRP. Furthermore, a participant questionnaire could be used to examine how participants perceive different levels of ambient noise and whether they believe the ambient noise affected their performance on the VRP.
REFERENCES


APPENDIX A

SUMMARY OF PRELIMINARY AND ACTUAL AMBIENT NOISE MEASUREMENTS
Summary of Preliminary and Actual Ambient Noise Measurements

An ideal environment for elicitation has ambient noise 10 dB lower than the quietest signal (Šrámková et al., 2015). When background noise is fewer than 10 dB below the signal, it becomes more likely that the noise is contributing to the sound level meter reading (Figure 15). Therefore, some studies recommend room noise at or below 40 dB A for VRP elicitation (Schutte & Seidner, 1983; Sulter et al., 1995). Although mouth to microphone distance may be decreased to account for background noise, mouth to microphone distance can only be decreased to a certain distance before near-field effects will render the signal distorted. This means that the signal can no longer be estimated accurately because its reverberation patterns are not comparable to those present at a greater distance. Although 40 dB A has been used as a maximum level for room noise, a recent study suggests that females can produce minimum-intensity levels between 41 and 53 dB A or 48 and 61 dB C (Šrámková et al., 2015). This study also found that males can produce minimum-intensity levels between 35 and 53 dB A or 49 and 64 dB C. Based on these results, the researchers recommend a maximum ambient noise level of 25 dB A during VRP elicitation.

Researchers have used dB A as well as dB C sound level meter settings for VRP elicitation. Sound level meter measurements made using a dB A setting filter out some low-frequency ambient noise, which is often emitted by electronic devices in the environment. Measurements made using a dB C setting filter out much less of this low-frequency noise (Figure 16). Although A weighting will filter out unwanted low-frequency background noise, it will also
have some effect on low-frequency signals. However, A weighting more closely mimics how humans perceive differences in sound intensity.

Figure 15: Background noise correction. A graph illustrating how many decibels should be subtracted from the total reading for specific difference between total noise and background noise.
In order to determine which sound level meter weighting (dB C or dB A) to use in this study, preliminary measurements of ambient noise in various environments were collected (Table 6). These environments were chosen as convenient and realistic places for VRP elicitation.
Table 6: Preliminary measurements of ambient noise in various environments using dB C and dB A weighting.

<table>
<thead>
<tr>
<th>Location</th>
<th>dB C</th>
<th>dB A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice rooms (quiet)</td>
<td>44-52 dB C</td>
<td>28-31 dB A</td>
</tr>
<tr>
<td>Practice rooms (piano &amp; violin playing down the hall)</td>
<td>48-54 dB C</td>
<td>31-38 dB A</td>
</tr>
<tr>
<td>Home voice studio</td>
<td>42-47 dB C</td>
<td>28 dB A</td>
</tr>
<tr>
<td>Clinic room (on a busy day)</td>
<td>59-62 dB C</td>
<td>33-34 dB A</td>
</tr>
</tbody>
</table>

Because this study seeks to standardize a practical method for VRP elicitation in convenient and realistic environments, A weighting was chosen. Preliminary measurements in realistic environments revealed that A-weighting would be more appropriate for VRP elicitation in these settings because the A-weighted measurements of ambient noise were lower and included less background noise. Furthermore, voice quality is a perceptual measure, and A weighting more closely mimics humans’ perception of intensity differences.

Although a study by Šrámková et al. (2015) suggested that background noise levels should not exceed 25 dB A, the current study recommends that background noise levels not exceed 40 dB A. Average background noise levels in realistic elicitation environments were between 28 and 38 dB A. In addition, the clinical implications of two or three additional decibels gained with lower background noise on the soft-intensity contour are unknown. The VRP has maintained its clinical utility despite background noise recommendations more lenient than those proposed by Šrámková et al. (2015). Because this study seeks to implement a clinician-controlled VRP protocol, clinical judgment should be used to elicit quiet productions during periods of decreased background noise. Furthermore, research suggests that normal
within-participant variation is approximately 2 dB (Coleman, 1993). Until further research is conducted to determine the clinically useful implications of a few additional decibels on the soft-intensity contour, this information will primarily be useful for researchers. Therefore, the 25 dB A criteria suggested by Šrámková et al. (2015) was deemed inappropriate for this protocol.

Ambient noise measurements were recorded using a sound level meter with A weighting during each elicitation session. Actual background noise for this study ranged from 28 dB A to 43 dB A (Table 7). Elicitation was conducted in practice rooms at two different universities, a music classroom, a music office, a church, a choir rehearsal room, and a clinic room.

<table>
<thead>
<tr>
<th>Location</th>
<th>Average ambient noise readings in dB A</th>
<th>Range of ambient noise in dB A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Practice rooms (at Northern Illinois University)</td>
<td>40</td>
<td>39-41</td>
</tr>
<tr>
<td>Practice rooms (at Saint Xavier University)</td>
<td>28.5</td>
<td>28-29</td>
</tr>
<tr>
<td>Music classroom</td>
<td>32.6</td>
<td>30-40</td>
</tr>
<tr>
<td>Music office</td>
<td>40.6</td>
<td>40-43</td>
</tr>
<tr>
<td>Church</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Clinic room</td>
<td>31.6</td>
<td>31-34</td>
</tr>
<tr>
<td>Choir rehearsal room</td>
<td>30</td>
<td>29-31</td>
</tr>
</tbody>
</table>
APPENDIX B

VHI-30
VHI-30

Prior to VRP elicitation, participants completed the Voice Handicap Index (VHI). A score above 28 points indicated a mild to severe voice disorder. Participants who scored above 28 points on the VHI were included in the study.

**VOICE HANDICAP INDEX**


These are statements that many people have used to describe their voices and the effects of their voices on their lives. Circle the response that indicates how frequently you have the same experience.

0 - never 1 - almost never 2 - sometimes 3 - almost always 4 - always

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My voice makes it difficult for people to hear me.</td>
</tr>
<tr>
<td>2</td>
<td>I run out of air when I talk.</td>
</tr>
<tr>
<td>3</td>
<td>People have difficulty understanding me in a noisy room.</td>
</tr>
<tr>
<td>4</td>
<td>The sound of my voice varies throughout the day.</td>
</tr>
<tr>
<td>5</td>
<td>My family has difficulty hearing me when I call them throughout the house.</td>
</tr>
<tr>
<td>6</td>
<td>I use the phone less often than I would like to.</td>
</tr>
<tr>
<td>7</td>
<td>I am tense when talking to others because of my voice.</td>
</tr>
<tr>
<td>8</td>
<td>I tend to avoid groups of people because of my voice.</td>
</tr>
<tr>
<td>9</td>
<td>People seem irritated with my voice.</td>
</tr>
<tr>
<td>10</td>
<td>People ask, “What’s wrong with your voice?”</td>
</tr>
<tr>
<td>11</td>
<td>I speak with friends, neighbors, or relatives less often because of my voice.</td>
</tr>
<tr>
<td>12</td>
<td>People ask me to repeat myself when speaking face-to-face.</td>
</tr>
<tr>
<td>13</td>
<td>My voice sounds raspy and dry.</td>
</tr>
<tr>
<td>14</td>
<td>I feel as though I have to strain to produce voice.</td>
</tr>
<tr>
<td>15</td>
<td>I find other people don’t understand my voice problem.</td>
</tr>
<tr>
<td>16</td>
<td>My voice difficulties restrict personal and social life.</td>
</tr>
<tr>
<td>17</td>
<td>The clarity of my voice is unpredictable.</td>
</tr>
<tr>
<td>18</td>
<td>I try to change my voice to sound different.</td>
</tr>
<tr>
<td>19</td>
<td>I feel left out of conversations because of my voice.</td>
</tr>
<tr>
<td>20</td>
<td>I use a great deal of effort to speak.</td>
</tr>
<tr>
<td>21</td>
<td>My voice is worse in the evening.</td>
</tr>
<tr>
<td>22</td>
<td>My voice problem causes me to lose income.</td>
</tr>
<tr>
<td>23</td>
<td>My voice problem upsets me.</td>
</tr>
<tr>
<td>24</td>
<td>I am less outgoing because of my voice problem.</td>
</tr>
<tr>
<td>25</td>
<td>My voice makes me feel handicapped.</td>
</tr>
<tr>
<td>26</td>
<td>My voice “gives out” on me in the middle of speaking.</td>
</tr>
<tr>
<td>27</td>
<td>I feel annoyed when people ask me to repeat.</td>
</tr>
<tr>
<td>28</td>
<td>I feel embarrassed when people ask me to repeat.</td>
</tr>
<tr>
<td>29</td>
<td>My voice makes me feel incompetent.</td>
</tr>
<tr>
<td>30</td>
<td>I am ashamed of my voice problem.</td>
</tr>
<tr>
<td>31</td>
<td>My voice is.</td>
</tr>
</tbody>
</table>
Participant Intake

All participants completed the following intake form prior to VRP elicitation. The form served as a screening measure to determine level of experience, vocal behaviors, and vocal hygiene behaviors.

Number: ____ Age: _______ Gender ______
Estimated daily voice use (talking):____________________________
Estimated daily voice use (singing):____________________________
Daily Fluid intake: ____________________________
Type of Fluids: (water, juice, coffee etc . . )__________________________
Tobacco use: ________________________
Number of years of classical singing training: _______________________
Reported singing range (soprano, mezzo-soprano, alto, tenor, baritone, bass, other)
Reported amount and type of performance: ________________________
Style of music (classical, opera, music theater, jazz, gospel, rock, pop, country, other)

Voice disorder diagnosis: ______________________________
Date of onset/duration: ______________________________
Medical treatment: ______________________________
Surgical treatment: ______________________________
Behavioral treatment: ______________________________
APPENDIX D

ELICITATION INSTRUCTIONS
Elicitation Instructions

Voice Range Profile Full Protocol

Setup
1. Have the person face you for modeling and monitoring of mouth to microphone distance
2. Distance from person’s mouth to the microphone of the sound level meter (SLM) should be 30 cm
3. Room acoustics should be at or below 40 dB A
4. Set the SLM to A weighting

Modeling and Cueing
1. Verbal encouragement refers to comments such as ‘great job, try to be that quiet again,’
   ‘good, try to be even quieter,’ etc.
2. Verbal feedback includes instructing the participant to use a different register (use your
   ‘head voice’ or ‘chest voice’) or instructing the participant to maintain the vowel, change
   mouth opening, move from a whisper to barely phonating, or moving from piano to a
   quieter production.
3. Imagery refers to verbal or nonverbal cues such as, ‘think about going up and over the
   pitch’ or using hand gestures to indicate descending or ascending loudness.

Training
1. What is a Voice Range Profile?
   a. Today, we want to find your voice range profile. This is a graph of how
      loudly and how softly you can sing every note in your range. (Show a sample
      VRP) The values across the bottom of the graph show which notes you sang.
      (Point to x-axis) The curve on the top will show how loud you were on each
      note (point to top curve), and the curve on the bottom will show how quiet
      you could be on each note (point to bottom curve). Every person has their
      own unique shape, so when we are done, we will have a record of your
      unique voice.
   b. Even though this is a measure of your singing voice, we want to measure all
      notes that you can produce. This means that we want to measure even the
      sounds that are not beautiful and that you would not want to sing in public.
2. How can we get your Voice Range Profile?
   a. Even though this is a measure of your singing voice, we want to measure all
      notes that you can produce. This means that we want to measure even the
      sounds that are not beautiful and that you would not want to sing in public.
   b. First, we will do some short warmups. Then, we will have you sing “ah” on
      using an exercise to get your range. Once we have your pitch range, we will
      have you sing “ah” as quietly as you can on different pitches in your range.
      We will use each pitch at least 3 times for both the loud and the soft. Then
we will have you sing “ah” as loudly as you can on each pitch. Again, you
will have at least 3 tries for each pitch. You will need to hold each pitch for
about 2 seconds so that I can record how loud it was. Feel free to ask
questions at any time. Do you have any questions right now? Ready to
warm up?

Warmups (adapted from NCVS, Titze)

1. Brief stretching (shoulder rolls, neck stretches, etc.)
2. Pitch glides on /u/
3. Semi-occluded vocal tract exercise to engage breath with vocal folds

Elicitation

1. Now we are going to find your pitch range. I will ask you to sing like this
   (demonstrate 8, 5, 3, 1, 3, 5, 8, 5, 3, 1 exercise on /a/), and we will first find your
   highest note, then we will find your lowest note. Remember, we are not looking for
   pretty sounds.
2. Pitch range
   a. First, obtain the pitch range by beginning on A₃ for females or A₄ for males
   b. Using an 8, 5, 3, 1, 3, 5, 8, 5, 3, 1 major scale exercise, move up from A by
      semitones to obtain the highest note in the person’s range. Record this note.
   c. Using the 8, 5, 3, 1, 3, 5, 8, 5, 3, 1 major scale exercise, move down from A by
      semitones to obtain the lowest note in the person’s range. Record this note.
3. Intensity range
   a. ***Things to remember: You are not looking for pretty sounds; you just want the
      loudest or softest sounds that the person can produce, so vibrato is fine. Allow
      the person time to breathe between pitches. Provide imagery (e.g., hand motion
to depict crescendo or decrescendo) on every trial. Recalibrate mouth to
      microphone distance after every trial.***
   b. Now we are going to see how quiet you can be on some notes in your pitch
      range. Remember, we are looking for very soft sounds. I will play a note,
      and I want you to sing /a/ on that note as quietly as you can. I need you to
      sing every note for at least 2 seconds. (demonstrate on example pitch)
      i. Start one perfect fifth above the lowest note obtained in the pitch range
         and have the person produce /a/ as quietly as he or she can on this note.
         Record the lowest number on the SLM for this semitone.
      ii. Next, provide a model, coaching, or feedback (as described at the
         beginning of this protocol).
      iii. Continue additional productions (up to a maximum of 6 productions) with
         coaching or cueing to obtain the 2 lowest productions within 2 dB of one
         another.
      iv. Proceed downward from this note to find the minimum-intensity for the
         lower range. Then, proceed upward from this note to find the minimum-
         intensity of semitones in the person’s upper range. Make sure you repeat
each semitone at least two times and record the intensity values for every semitone.

v. Recalibrate mouth to microphone distance after every trial.

c. Now we are going to see how loud you can be on some notes in your pitch range. Remember, we are looking for very loud sounds. I will play a note, and I want you to sing /a/ on that note as loudly as you can. I need you to sing every note for at least 2 seconds. (demonstrate on example pitch)

i. Start one perfect fifth above the lowest note obtained in the pitch range and have the person produce /a/ as loudly as he or she can on this note. Record the highest number on the SLM for this semitone.

ii. Next, provide a model, coaching, or feedback (as described at the beginning of this protocol).

iii. Continue additional productions (up to a maximum of 6 productions) with coaching or cueing to obtain the 2 highest productions within 2 dB of one another.

iv. Proceed downward from this note to find the maximum-intensity for the lower range. Then, proceed upward from this note to find the maximum-intensity of semitones in the person’s upper range. Make sure you repeat each semitone at least two times and record the intensity values for every semitone.

v. Recalibrate mouth to microphone distance after every trial.

Voice Range Profile Shortened Protocol

Setup
1. Have the person face you for modeling and monitoring of mouth to microphone distance
2. Distance from person’s mouth to the microphone of the SLM should be 30 cm
3. Room acoustics should be at or below 40 dB A
4. Set the SLM to A weighting
5. If you will be obtaining a voice range profile again for this person, do so at the same time of day.

Modeling and Cueing
1. Verbal encouragement refers to comments such as ‘great job, try to be that quiet again,’ ‘good, try to be even quieter,’ etc.
2. Verbal feedback includes instructing the participant to use a different register (use your ‘head voice’ or ‘chest voice’) or instructing the participant to maintain the vowel, change mouth opening, move from a whisper to barely phonating, or moving from piano to a quieter production.
3. Imagery refers to verbal or nonverbal cues such as, ‘think about going up and over the pitch’ or using hand gestures to indicate descending or ascending loudness.
Training

1. **What is a Voice Range Profile?**
   a. **Today, we want to find your voice range profile.** This is a graph of how loudly and how softly you can sing on notes in your range. *(Show a sample VRP)* The values across the bottom of the graph show which notes you sang. *(Point to x-axis)* The curve on the top will show how loud you were on each note *(point to top curve)*, and the curve on the bottom will show how quiet you could be on each note *(point to bottom curve)*. Every person has their own unique shape, so when we are done, we will have a record of your unique voice.

2. **How can we get your Voice Range Profile?**
   a. **Even though this is a measure of your singing voice, we want to measure even the sounds that are not beautiful and that you would not want to sing in public.**
   b. **First, we will do some short warmups.** Then, we will have you sing “ah” on using an exercise to get your range. Once we have your pitch range, we will have you sing “ah” as quietly as you can on different pitches in your range. We will use each pitch at least 3 times for both the loud and the soft. Then we will have you sing “ah” as loudly as you can on each pitch. Again, you will have at least 3 tries for each pitch. You will need to hold each pitch for about 2 seconds so that I can record how loud it was. Feel free to ask questions at any time. Do you have any questions right now? Ready to warm up?

**Warmups (adapted from NCVS, Titze)**

4. Brief stretching (shoulder rolls, neck stretches, etc.)
5. Pitch glides on /u/
6. Semi-occluded vocal tract exercise to engage breath with vocal folds

**Elicitation**

1. Use pitch range obtained in full protocol with additional sampling at extremes if the participant is able to produce additional semitones.
2. **Intensity range**
   a. ***Things to remember: You are not looking for pretty sounds; you just want the loudest or softest sounds that the person can produce, so vibrato is fine. Allow the person time to breathe between pitches. Provide imagery *(e.g., hand motion to depict crescendo or decrescendo)* on every trial. Recalibrate mouth to microphone distance after every trial.***
   b. For this protocol, choose notes that correspond with every perfect fifth and every octave starting from the minimum pitch in the participant’s pitch range *(e.g. A minimum of C₂ means you will be using C₂, G₂, C₃, G₃, etc.).* You may sample additional points at the extremes of the person’s range.
i. If there is a more than 7 dB difference between sampling points, sample
between these points. Continue to sample between the points if the
difference in dB between points is greater than 7 dB.

c. Now we are going to see how quiet you can be on some notes in your pitch
range. Remember, we are looking for very soft sounds. I will play a note,
and I want you to sing /a/ on that note as quietly as you can. I need you to
sing every note for at least 2 seconds. (demonstrate on example pitch)
   i. Start one perfect fifth above the lowest note obtained in the pitch range
   and have the person produce /a/ as quietly as he or she can on this note.
   Record the lowest number on the SLM for this semitone.
   ii. Next, provide a model, coaching, or feedback (as described at the
   beginning of this protocol).
   iii. Continue additional productions (up to a maximum of 6 productions) with
   coaching or cueing to obtain the 2 lowest productions within 2 dB of one
   another.
   iv. Proceed downward (using perfect fifth and octave method with additional
   sampling if needed) from this note to find the minimum-intensity for the
   lower range. Then, proceed upward from this note to find the minimum-
   intensity of semitones in the person’s upper range. Make sure you repeat
each semitone at least two times and record the intensity values for every
semitone.
   v. Recalibrate mouth to microphone distance after every trial.

d. Now we are going to see how loud you can be on some notes in your pitch
range. Remember, we are looking for very loud sounds. I will play a note,
and I want you to sing /a/ on that note as loudly as you can. I need you to
sing every note for at least 2 seconds. (demonstrate on example pitch)
   i. Start one perfect fifth above the lowest note obtained in the pitch range
   and have the person produce /a/ as loudly as he or she can on this note.
   Record the highest number on the SLM for this semitone.
   ii. Next, provide a model, coaching, or feedback (as described at the
   beginning of this protocol).
   iii. Continue additional productions (up to a maximum of 6 productions) with
   coaching or cueing to obtain the 2 highest productions within 2 dB of one
   another.
   iv. Proceed downward (using perfect fifth and octave method with additional
   sampling if needed) from this note to find the minimum-intensity for the
   lower range. Then, proceed upward from this note to find the maximum-
   intensity of semitones in the person’s upper range. Make sure you repeat
each semitone at least two times and record the intensity values for every
semitone.
   v. Recalibrate mouth to microphone distance after every trial.