Repeated Measures of Vocal Fundamental Frequency Perturbation Obtained Using the Visi-Pitch

Amy Dwire and Rebecca McCauley

Department of Communication Sciences, University of Vermont, Burlington, Vermont, U.S.A.

Summary: This article reports on a study of intrasubject variability for a measure of vocal fundamental frequency perturbation in a group of young, normal speakers. Measurements of relative average perturbation (RAP) obtained using a Visi-Pitch were examined for the vowels /a/, /i/, and /u/ produced by 25 women and 24 men on two occasions, ~1 week apart. Important findings consisted of higher levels of RAP for women than men and greater variability of RAP over the time period studied for women than men. Conclusions include the need for additional normative data regarding the use of the Visi-Pitch for the purpose of examining jitter. Key Words: Jitter--Gender--Norms--Variability.

Measures of cycle-to-cycle variation in fundamental frequency, or jitter, are of interest as an objective means of differentiating normal from disordered voices and of quantifying the severity of vocal deviation (1). For normal speakers, contributors to jitter are thought to consist of a wide range of factors including variability in vocal fold mass, distribution of mucus on the folds, symmetry of vocal fold structure, timing of laryngeal muscle action potentials, and glottal flow characteristics (2,3). For speakers with abnormal voice quality, measures of jitter are thought to reflect those same factors as well as others that may be unique to the processes or structural changes underlying the disorder (4). The potential explanatory role of jitter and the fact that jitter provides a noninvasive estimate of laryngeal function make it of great potential utility in the clinic (5). To date, however, methodological complexities and insufficient normative data impede the use of jitter for clinical purposes.

Methodologically, researchers studying vocal fundamental frequency perturbation have used a diversity of instrumental means for determining the duration of each glottal period and an even greater diversity of statistical methods for summarizing patterns of durational differences across successive periods (6,7). Because of these varied methodologies, the comparability of data obtained in different studies has been difficult to determine and associated differences in findings difficult to reconcile (1).

At least in part due to the range of methods used, the diagnostic validity of vocal fundamental frequency perturbation measures has not been clearly established. Although several researchers have demonstrated the utility of such measures in the detection of laryngeal pathology (8,9), others have deemed them lacking in both sensitivity and specificity—that is, in the ability to correctly identify true positives while at the same time excluding false positives (10,11).

Of particular practical significance, however, is the absence of data concerning the intrasubject variability of clinically available measures of jitter, such as those obtained using the Visi-Pitch—a tool widely used by speech-language pathologists (7). In the manual of the Visi-Pitch (12), only rough guidelines are given for normal perturbation values for
men and women. Specifically, the manual states that normal values of pitch perturbation using Koike's formula are $\sim 1.00\%$ and below. However, no other normative data are provided. Additional information concerning such measures is needed before clinicians can confidently use clinical measures of jitter to draw inferences about laryngeal function, at initial diagnosis of voice abnormality or during the course of treatment. The purpose of this study, therefore, was to obtain information about the statistical stability of a readily available clinical measure of jitter for a group of young adults with normal voice quality who were tested twice, in sessions $\sim 1$ week apart.

**METHODS**

**Subjects**

Forty-nine subjects participated in this study: 24 men and 25 women ranging in age from 18 to 25 years (overall mean = 20.8 years). Subjects in this age range were used because of their availability and because the relatively high incidence of certain laryngeal pathologies in that age group (i.e., vocal polyps and edema) suggests their value as a normative population (13). Potential participants were asked to fill out a questionnaire to determine eligibility. Subjects were considered suitable if they were nonsmokers who had no personal or family history of laryngeal pathology or voice problems, as self-reported. They were also required not to have had an upper respiratory tract infection, sinus infection, or laryngitis within 2 weeks of entering the study. Prior to their participation, all subjects were judged to be free of abnormal voice quality by one of the experimenters.

**Equipment**

The Visi-Pitch microprocessor (Kay Elemetrics model 6087PC) was used to measure vocal fundamental frequency perturbation. With use of this system, acoustic input signals were subjected to low-pass filtering (with cutoff frequencies resulting in a 3-dB decrement at 100 and 313 Hz for male and female speakers, respectively), then analog-to-digital conversion at a resolution of 14 bits. The sampling rate reported in the Visi-Pitch manual is 100 kHz. The system uses a peak-picking method and a modified version of Koike's formula (14) to calculate relative average perturbation (RAP; %), a moving average measure of cycle-to-cycle changes in vocal fold vibration.

Prior to the experiment, the Visi-Pitch microprocessor was examined to assess the instrument's internal jitter. A programmable pulse/function generator (Hewlett Packard model 8116A) was used to produce the pure tone signals of 100 and 200 Hz that were used in the calibration. This generator possesses negligible internal jitter of 0.03% root mean square. With the Visi-Pitch, jitter was calculated as essentially zero. These readings suggested that the Visi-Pitch itself did not contribute to observed jitter values because of internal noise.

**Sampling and acoustic analysis procedures**

Each subject who participated in the study was seen individually for two recording sessions, each lasting $\sim 20$ min. In the initial session, the subject was required to fill out a questionnaire and was then oriented to the task and asked to use his/her habitual pitch and intensity, which were described as those of an everyday, "comfortable" voice. Following that introduction, each subject participated in 15 experimental trials. An additional 15 trials were produced in a second session conducted no earlier than 7 and no later than 14 days following the first session. Two subjects had upper respiratory infections at the appointed time for their second session 1 week later; they were instructed to return on a date that was no later than 2 weeks after their first session. Both were judged by one of the experimenters to have normal voice quality at that time and their performances were not associated with abnormally high jitter values.

For all trials, the subject was instructed to hold the microphone $\sim 2.5-3.5$ cm from the mouth (the distance recommended by the Visi-Pitch manual) when phonating. That instruction, rather than the establishment of a single, constant distance, was used to reproduce commonly occurring clinical conditions. Each subject was asked to produce 15 experimental trials consisting of five 6-s repetitions of three vowels: /a/, /i/, and /u/. This vowel order was chosen arbitrarily and kept constant across subjects to produce results that can be generalized directly to data obtained using the same standard procedure.

All Visi-Pitch analyses were performed on live productions because of concerns regarding tape recorder effects on jitter (15). After each trial (i.e., one repetition of one vowel) was captured using the Visi-Pitch, the original 6-s sample was restricted to
A 4-s analysis sample using cursors provided by the software. This procedure was used to exclude the onset and offset of each vocal production. Because a constant time interval (rather than a specific number of cycles) was used, the number of cycles analyzed for each individual ranged from means of $\sim 511$ (128 Hz $\times$ 4 s) for male subjects to 933 (233 Hz $\times$ 4 s) for female subjects. For each production, mean perturbation (RAP) and mean fundamental frequency ($F_0$) data were calculated on the Visi-Pitch and then transferred to a data sheet manually for later statistical analysis.

**Statistical analysis procedures**

All analyses were performed using JMP version 2.0.4 (16). Two methods were used to examine the statistical stability of RAP across the two testing sessions. The first method consisted of an analysis of variance performed on log mean RAP as the dependent measure. Mean data were initially obtained by summing across the five productions obtained from each subject for each vowel within a session. A logarithmic transformation was applied to the resulting data because initial examination of the distributions obtained for subgrouping by gender and vowel (e.g., female productions of /i/) suggested that the data were not normally distributed (17). The transformation resulted in data that appeared normal in distribution. The analysis of variance on log mean RAP used a repeated measures design, with gender as a between-subject factor and session and vowel as within-subject repeated measures. This approach was used to examine the overall effects of vowel and session on the magnitude of RAP scores for the male and female subject groups. Although log-transformed data were used in these analyses, untransformed mean data will be reported for the sake of clarity.

As a second method for examining the temporal stability of RAP values obtained on the Visi-Pitch, Pearson product–moment correlation coefficients were calculated between RAP scores for sessions 1 and 2. This approach was used to examine the consistency of RAP scores for individual subjects across a time period analogous to that between weekly treatment sessions.

As an additional means of examining intrasubject variability for RAP, coefficient of variation (CV; i.e., SD/$\bar{x}$) was obtained for each subject on each vowel within a session. The resulting values were then compared with a cutoff value of 30%, a level that has been recommended as the highest level of variability that should be regarded acceptable for clinical measures of voice (18).

A fourth and final analysis consisted of a comparison of the RAP obtained for each subject against the cutoff for identification of normal RAP values recommended for the Visi-Pitch (12). Because a sample of speakers with normal voice quality and normal history had been sought, it was expected that measures below the cutoff value would predominate. Scores higher than that value were considered false positives.

**RESULTS AND DISCUSSION**

Findings regarding the overall levels of RAP obtained in these groups of young normal women and men, as well as findings regarding statistical stability of RAP between and within each session, are discussed in terms of the research literature and clinical implications.

**Overall levels of RAP and between-session stability**

RAP and $F_0$ for females and males are reported in Table 1. The analysis of variance used to examine the effects of gender, vowel, and session on log mean RAP produced significant main effects of gender [$F(1,47) = 13.4177$, $p < 0.0006$] and vowel [$F(2,47) = 15.4529$, $p < 0.0000$], as well as a significant interaction between gender and vowel [$F(2,235) = 4.5732$, $p < 0.0113$]. Post hoc comparisons demonstrated significantly higher RAP values for women than men for each vowel ($p < 0.0000$ for each comparison) and differing patterns of vowel RAP values for men versus women (see Fig. 1). Specifically, for the female subjects, RAP was highest for /a/ ($\bar{x} = 0.891\%$), next highest for /u/ ($\bar{x} = 0.840\%$), and lowest for /i/ ($\bar{x} = 0.536\%$); whereas for male subjects, the only significant difference between vowels was that between /u/ ($\bar{x} = 0.583\%$) and /i/ (0.420%). The greater magnitude of jitter for /u/ versus /i/ differs from Horii's finding of no significant differences among vowels for jitter resulting from his study of an exclusively male sample (19). The lack of significant effects involving session failed to document systematic changes in RAP over the 1-week time period.

Between-session stability of RAP was also examined through the calculation of correlation coefficients for session 1 versus session 2 RAP values (see Table 2). The correlations were almost all sta-
TABLE 1. Means (SD) for relative average perturbation (RAP; %) in session 1 and session 2 and combined data and for overall fundamental frequency (Hz)

<table>
<thead>
<tr>
<th></th>
<th>Females</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/a/</td>
<td>/i/</td>
<td>/u/</td>
<td>/a/</td>
<td>/i/</td>
<td>/u/</td>
</tr>
<tr>
<td>RAP for</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>session 1</td>
<td>1.044 (0.735)</td>
<td>0.606 (0.620)</td>
<td>0.961 (1.280)</td>
<td>0.394 (0.192)</td>
<td>0.439 (0.688)</td>
<td>0.682 (1.185)</td>
</tr>
<tr>
<td>RAP for</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>session 2</td>
<td>0.738 (0.448)</td>
<td>0.467 (0.209)</td>
<td>0.719 (0.524)</td>
<td>0.375 (0.146)</td>
<td>0.401 (0.421)</td>
<td>0.484 (0.380)</td>
</tr>
<tr>
<td>RAP for</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>both sessions</td>
<td>0.891 (0.622)</td>
<td>0.536 (0.463)</td>
<td>0.840 (0.976)</td>
<td>0.385 (0.169)</td>
<td>0.420 (0.565)</td>
<td>0.583 (0.877)</td>
</tr>
<tr>
<td>F0 for</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>both sessions</td>
<td>222.92 (20.25)</td>
<td>234.73 (18.85)</td>
<td>241.85 (22.46)</td>
<td>117.77 (15.51)</td>
<td>128.12 (26.52)</td>
<td>137.18 (44.36)</td>
</tr>
</tbody>
</table>

A comparison of the absolute and relative jitter values obtained here for men versus women with those reported in the literature is complicated by differences in methodology. A study by Stone and Rainey (18) provides the closest match to the present study in terms of methodology, yet it included only women subjects. Similarities in the methodology of the two studies included the use of the same measure of jitter (RAP) obtained on a closely related instrument (a different model of Visi-Pitch) for sustained productions of the same three vowels. Specific findings of Stone and Rainey consisted of mean RAP values of 0.826% for /a/, 0.861% for /i/, and 0.987% for /u/ compared with 0.891% for /a/, 0.536% for /i/, and 0.840% for /u/ for the female subjects described here. This comparison reveals relatively consistent magnitudes across studies for /a/ and /u/, but not /i/.

No study to date has examined the RAP values for men versus women using the Visi-Pitch. However, in a study in which RAP was obtained using different instrumentation, higher RAP values for men than women were obtained (20). Similarly, when comparisons are made with studies that used other measures of jitter, the present finding of higher levels of jitter for women than men is at odds with findings suggesting the reverse (3,11,21, 22) as well as with findings suggesting relatively equivalent RAP values for men and women speaking at their preferred pitch (23).

Other studies in which between-session stability has been examined include those of Stone and Rainey (18) and Higgins and Saxman (3). RAP was compared for eight women in Stone and Rainey's study across an 8-week interval. Resulting correlations were relatively high only for /i/ and /a/ (viz., 0.707 with p = 0.025 for /i/ and 0.775 with p = 0.012 for /a/). Higgins and Saxman examined jitter factor in 10 women and 5 men at 2-day intervals across 33 days. In those women, significant changes in jitter were associated with changing ovarian hormone fluctuations across the menstrual cycle. In fact, al-
though Higgins and Saxman noted relatively consistent jitter values during premenstrual and menstrual intervals, relatively large changes in jitter were noted at ovulation. Their findings provide a reasonable explanation of the high degree of variability observed in the present group of young women whose vocal behavior was studied over a 1-week interval.

**Within-session stability**

When the mean CV was calculated for each group and vowel within each session, it was found that only two values for the women versus five values for the men fell below the 30% of variation in the mean recommended by Stone and Rainey (18) (see Table 3). For women, CV ranged from 26.87 to 35.60% and for men from 24.24 to 32.06%. Further, only /a/ and /u/ for the male subjects fell below the criterion for both sessions. The values obtained in this study for women subjects compare favorably with those reported by Stone and Rainey for their female subjects, for whom CV values ranged from 30% for /i/ to 48% for /a/.

**Comparison of RAP values with Visi-Pitch guidelines**

Figure 2 for female subjects and Fig. 3 for male subjects illustrate the comparison of RAP values for each vowel and session against the 1.00% cutoff recommended by the Visi-Pitch manual. As shown in Fig. 2, 24 of the women’s 150 RAP values exceeded that value. In addition, 13 women, or 52% of those sampled, would have been identified as demonstrating abnormally high levels of jitter on one or more vowels using these methods. Because normal

### TABLE 2. Pearson product–moment correlation coefficients between relative average perturbation for session 1 and 2

<table>
<thead>
<tr>
<th>Gender</th>
<th>Session</th>
<th>/a/</th>
<th>/i/</th>
<th>/u/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Females (n = 25)</td>
<td></td>
<td>0.58</td>
<td>0.36</td>
<td>0.42</td>
</tr>
<tr>
<td>Males (n = 24)</td>
<td></td>
<td>0.88</td>
<td>0.83</td>
<td>0.90</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vowels</th>
<th>Females (n = 25)</th>
<th>Males (n = 24)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/a/</td>
<td>/i/</td>
</tr>
<tr>
<td>r value</td>
<td>0.002</td>
<td>0.079</td>
</tr>
<tr>
<td>p value</td>
<td>0.002</td>
<td>0.079</td>
</tr>
</tbody>
</table>

### TABLE 3. Mean coefficient of variation for relative average perturbation values

<table>
<thead>
<tr>
<th>Gender</th>
<th>Session</th>
<th>/a/</th>
<th>/i/</th>
<th>/u/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female subjects (n = 25)</td>
<td></td>
<td>35.60</td>
<td>32.03</td>
<td>30.87</td>
</tr>
<tr>
<td>1</td>
<td>30.78</td>
<td>26.87</td>
<td>28.88</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>25.97</td>
<td>32.06</td>
<td>29.71</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>24.24</td>
<td>27.72</td>
<td>27.56</td>
<td></td>
</tr>
</tbody>
</table>

**FIG. 2.** Comparison of relative average perturbation values with recommended cutoff of 1.0% for female subjects.
subjects had been used in the present study, this finding can be interpreted as suggesting a very high false-positive rate for others who might use the same procedures and cutoff (i.e., those outlined by the Visi-Pitch manual).

As shown in Fig. 3, only eight of the men’s 150 RAP values exceeded the recommended cutoff value of 1.0. In addition, only three of the male subjects, or 12.5% of that group, would have been identified as demonstrating abnormally high levels of jitter on one or more vowels. This false-positive rate appears appreciably better than that obtained for the female subjects in this study.

CONCLUSIONS

This study was undertaken to study the variability of a measure of jitter obtained using the clinically ubiquitous Visi-Pitch (Kay Elemetrics model 6087PC) for a group of young adults with normal voice quality. It was undertaken with hopes of providing information that might prove valuable for individuals contemplating the clinical use of that measure. Consequently, procedures used in the study were chosen to reflect those recommended by the manufacturer of that instrument and those likely to be followed in a typical clinical setting.

Findings included the following observations regarding the comparability of our findings with those obtained by other researchers: Levels of jitter for women were roughly comparable with those obtained for the one other study in which RAP has been used (18). However, the finding of higher levels of jitter for women than men contrasted with the results of most other comparative studies of jitter (albeit using measures other than RAP) that have reported higher or equivalent levels for men versus women (3,11,21,22). Although levels of jitter across a 1-week time period remained relatively consistent for groups of subjects, the relationship of between-session values for individual subjects was quite marked for men, but only moderate for women. Within-session intrasubject variability, as measured using the CV, suggested greater stability for men than women. And finally, comparisons of individual RAP values against a diagnostic cutoff revealed high false-positives for women and lower, more acceptable false-positive rates for men.

Clinical implications of these findings differ for the two groups of subjects, with RAP appearing more appropriate for men than women—both for purposes of identifying abnormal vocal performance and for tracking vocal performance over time. This conclusion is supported for diagnostic purposes by the measures of within-session stability (CV) that were within the recommended range for men but not women and by the results of comparisons against recommended cutoffs that showed higher false-positives for women than men. The conclusion that RAP is more appropriate for use with men than women is supported for tracking of vocal performance by the findings of group consis-
tency in RAP across sessions and of higher be-
tween-session correlations for men than women.

Several methodological limitations of this study,
however, must be taken into account when evalu-
ating both the clinical implications described herein
and the research implications to be drawn from the
present study. First, because of research suggesting
hormonal effects on jitter for women speakers (3),
future research and clinical practice should ac-
knowledge the possibility of these effects. A second
limitation exists in the selection of speakers with
normal voice as subjects. Clearly, extrapolations
from a measure characterizing variability of normal
speakers to that expected in speakers without nor-
mal voice quality are tentative at best (22). Addition-
ral research is needed to provide a firmer empir-
ical basis for such conjecture. Finally, by restricting
measurement to RAP obtained using the Visi-Pitch,
the ability to determine what variability was likely
to be due to real differences in vocal performance
as opposed to sources of error attributable to the Visi-
Pitch or procedures used with it was lost. Thus,
it appears that additional research is needed to help
clarify both the value of and suitable guidelines for
the use of this measure in everyday clinical prac-
tice.

Acknowledgment: We gratefully acknowledge the assis-
tance of Barry Guitar, Richard Absher, Gayle Belin, Jim
Lubker, and Chris Murphy who provided valuable com-
ments on drafts of this manuscript. Special thanks are
extended to Alan Howard for his helpfulness and incred-
ibly clear advice regarding the statistics.

REFERENCES

1. Karnell M, Scherer R, Fischer L. Comparison of acoustic
voice perturbation measures among three laboratories. J
2. Titze IR, Horii Y, Scherer RC. Some technical consider-
atations in voice perturbation measurements. J Speech
3. Higgins M, Saxman J. Variations in vocal frequency pertur-
4. Koike Y, Takahashi H, Calcutt TE. Acoustic measures for
detecting laryngeal pathology. Acta Otolaryngol 1977;84:
105-17.

5. Deem JF, Knack JV, Manning WH, Matesich J. The auto-
matic extraction of pitch perturbation using microcom-
puters: some methodological considerations. J Speech
6. Heiberger VL, Horii Y. Jitter and shimmer in sustained pho-
nation. In: Lass N, ed. Speech and language. New York:
Academic Press, 1982:292-322. (Advances in basic research
and practice; vol 7).
7. Baken RJ. Clinical measures of speech and voice. Boston:
8. Klingholz F, Martin F. Quantitative spectral evaluation of
9. Askensfelt A, Hammarberg B. Speech waveform perturba-
tion analysis: a perceptual acoustical comparison of seven
10. Zyski BJ, Bull GL, McDonald WE, Johns ME. Perturbation
analysis of normal and pathological larynges. Folia Phoniatr
11. Ludlow CL, Bassich CJ, Connor NP, Coulter DC, Lee YJ.
The validity of using jitter and shimmer to detect laryngeal
pathology. In: Baer T, Sasaki C, Harris K, eds. Laryngeal
function in phonation and respiration; vocal fold physiology
13. Herrington-Hall B, Lee L, Stemple J, Neime K, Miller-
McHone M. Description of laryngeal pathologies by age,
sex, and occupation in a treatment-seeking sample. J Speech
Hear Res 1988;31:57-64.
14. Koike Y. Application of some acoustic measures for the eval-
uation of laryngeal dysfunction. Studia Phonol 1973;7:
17-23.
15. Doherty ET, Shipp T. Tape recorder effects on jitter and
16. SAS Institute. JMP user's guide: version 2 of JMP. Cary,
17. Bartlett MS, Kendall DG. Statistical analysis of varianc-
eterogeneity and the logarithmic transformation. JRSS
1946;8:128-38.
18. Stone RE, Rainey CL. Intra- and intersubject variability in
19. Horii Y. Jitter and shimmer: differences among sustained
20. Till JA, Jafari M, Crumley RL, Law-Till CB. Effects of ini-
tial consonant, pneumotachographic mask, and oral pres-
sure tube on vocal perturbation, harmonics-to-noise, and in-
21. Casper JC. Frequency perturbation in normal speakers. A
descriptive methodological study. Doctoral dissertation,
Syracuse University, 1983.
22. Nittouer S, McGowan RS, Milenkovic PH, Beehler D.
Acoustic measurements of men's and women's voices: a study
of context effects and covariation. J Speech Hear Res
1990;33:761-75.
23. Orlikoff R, Baken R. Vocal jitter at different fundamental
frequencies. Paper presented at the 17th Symposium: Care