What is “Twang”?

*Johan Sundberg and †Margareta Thalén, *|Stockholm, Sweden

**Summary:** A single female professional vocal artist and pedagogue sang examples of “twang” and neutral voice quality, which a panel of experts classified, in almost complete agreement with the singer’s intentions. Subglottal pressure was measured as the oral pressure during the occlusion during the syllable /pæ/. This pressure tended to be higher in “twang,” whereas the sound pressure level (SPL) was invariably higher. Voice source properties and formant frequencies were analyzed by inverse filtering. In “twang,” as compared with neutral, the closed quotient was greater, the pulse amplitude and the fundamental were weaker, and the normalized amplitude tended to be lower, whereas formants 1 and 2 were higher and 3 and 5 were lower. The formant differences, which appeared to be the main cause of the SPL differences, were more important than the source differences for the perception of “twanginess.” As resonatory effects occur independently of the voice source, the formant frequencies in “twang” may reflect a vocal strategy that is advantageous from the point of view of vocal hygiene.

**Key Words:** Nonclassical singing–Voice source–Formant frequencies–Subglottal pressure.

**INTRODUCTION**

The term “twang” is frequently used in descriptions of vocal styles, particularly those used in some contemporary popular music, for example, pop, rock, country, and musical theater. It is typically associated with loud and high-pitched singing and is used to create an impression of energy and expressivity. According to some authors, “twang” is also used in male classical operatic singing, particularly tenor voices. Kayes (2000) claims that classically trained male singers, for example, Luciano Pavarotti, sing with “twang.” “Twang” is also used as a term for a specific vocal technique. For example, Estill describes a certain vocal technique as producing a loud, brassy, and sometimes twangy sound. Yanagisawa and Estill (1989) regard “twang” as a voice-quality category separate from, for example, opera, sob, and belting.

Considering the substantial timbral difference between operatic and popular music styles of singing, several different meanings appear to be attributed to the term “twang.” Voice research should be capable of resolving such a terminological confusion by describing the acoustic, perceptual, and physiological characteristics. On the basis of such data, a clearer definition of “twang” might be proposed.

Using fiberscope imaging, Yanagisawa and Estill (1989) analyzed the pharyngeal characteristics of the “twang” quality. They found that the aryepiglottic sphincter was narrowed in “twang” production. However, this trait was not specific to “twang,” because it was observed also in the production of opera and belt qualities.

In a synthesis experiment, Titze et al. (2003) attempted to identify the characteristics of “twang.” Their synthesis system was constituted by a glottal airflow model combined with vocal tract–area function model. The synthesis was a set of repetitions of the syllable /ya/ produced with a speech-like intonation. The open quotient of the voice source was varied along a continuum ranging from normal to “twang.” The vocal tract parameters that were varied were overall length and pharyngeal constriction. A listening panel rated the degree of “twang” perceived in these synthesized stimuli. The results showed that all three parameters affected the mean rating values. Thus, the perception of “twang” increased with decreasing open quotient, vocal tract length, and pharyngeal area.

The purpose of the present investigation was to complement the Yanagisawa and Estill and Titze et al. investigations with acoustic analyses of professional singing. Thus, recordings were made under well-controlled conditions of a professional singer producing contrasting examples of “twang” and neutral singing styles. The recorded material was analyzed in various ways with the aim to identify voice source properties and formant frequency combinations that contribute to listeners’ perception of “twang” quality.

**METHOD**

Recordings were made of a single female professional vocal artist and pedagogue, specialized in jazz and pop genres. Her task was to sing the pitches C4, E4, G4, C5 in “twang” style and in neutral style. She sang a diminuendo from ff to pp, repeating the syllable /pæ/ on each pitch. In addition, she sang a phrase of a song from Chaka Khan’s repertoire, “Ain’t nobody” by D. Wolinski (Figure 1). She sang this song two times with the original lyrics and two times with the syllable /pæ/ replacing the original lyrics.

The recordings were made in the anechoic chamber of the Department of Linguistics, Stockholm University. Four signals were digitized: (1) audio recorded by a condenser microphone (B&K 4003, Bruel & Kjaer, Naerum, Denmark), hanging from the ceiling at a distance of 1.5 m; (2) flow recorded by means of a pneumotachograph mask (MSFI 2; Glottal Enterprises, Syracuse, New York); (3) oral pressure recorded by means of a pressure transducer mounted at the end of a thin plastic tube which the subject held in the corner of her mouth; and (4) electroglottograph signal EGG (Glottal Enterprises). The recordings were directly digitized with a sampling frequency of 64 kHz and stored in computer files using the Swell Signal Workstation (Saven Hitech, Sverige, Taby, Sweden). Signal
calibrations were also recorded on the computer files. Sound level was calibrated by recording a B&K calibrator (Brüel & Kjaer, Naerum, Denmark), and the sound pressure level (SPL) of the calibration tone was announced on the computer file. Flow calibration was performed by means of the Glottal Enterprises flow calibration equipment. Oral pressure was calibrated by recording a set of pressure values determined by means of a manometer, and these values were recorded and announced in the computer file.

The audio signal was inverse filtered using the custom-made Decap program (Svante Granqvist), complemented by the derivative of the EGG waveform, dEGG. Two criteria were applied for tuning the inverse filters: (1) a ripple-free closed phase, (2) synchronicity of the main positive peak of the dEGG and the flow waveform discontinuity representing the closing of the glottis, and, whenever applicable, synchronicity of the negative peak of the dEGG and the flow waveform discontinuity representing the opening of the glottis. The resulting flow glottograms were saved to files together with their derivatives and the dEGG signals.

The following flow glottogram measures were determined (Figure 2): period $T_0$, closed-phase $Q_{\text{Closed}}$, AC flow amplitude, maximum flow declination rate (MFDR), and the level of the voice source fundamental relative to that of the second source spectrum partial, $H_1 - H_2$. $Q_{\text{Closed}}$ was calculated as the ratio between the closed-phase duration and $T_0$, and the normalized amplitude quotient (NAQ), previously found to be related to perceived degree of phonatory pressedness, as the ratio between the flow AC amplitude and the product of MFDR and $T_0$. Subglottal pressure was estimated from the oral pressure during the occlusion for the consonant /p/.

As only one subject was recorded in our experiment, it was important to find out how representative the recorded examples were of what was generally understood by the term “twang.” Therefore, a listening test was carried out with six experts, all singers and singing teachers who were well acquainted with “twang.”

A total of 44 long tones were copied from the recording of the /pae/ version of the song Ain’t nobody, 22 tones sung in neutral and 22 in “twang.” Using the custom-made program Glue (Svante Granqvist), the tones were arranged in six different random orders in six sound files that were transferred to CDs. In each file, each of the 44 stimuli appeared twice. Thus, the test material consisted of a total of 88 stimuli. The stimuli were separated by 3-second silent intervals. The duration of the entire test was 303 seconds.

Each panelist was given a CD together with an answer sheet. The listeners’ task was to rate the degree of “twanginess” on 100-mm visual analog scales with extremes labeled “not at all typical” and “very typical.”

![FIGURE 1](image1.png)

**FIGURE 1.** Beginning of the triad task in which the subject sang a diminuendo from ff to pp, repeating the syllable /pæ/ on each of the pitches C4, E4, G4, C5 in “twang” style and in neutral style.

![FIGURE 2](image2.png)

**FIGURE 2.** Parameters measured in the flow glottogram and its derivative.
RESULTS

Listening test

Listeners' consistency was analyzed in terms of the correlation in the ratings of “twanginess” for repeated stimuli. The results are shown in Table 1. A perfect correlation would yield a determination constant and a slope of 1.0 and an intercept of 0. The determination constant varied between 0.350 and 0.937, and the slope was close to 1 for all subjects except one (listener 4).

The mean values of the ratings collected in the listening test are shown together with the standard deviations in Figure 3. There is a clear division of the means, none of the stimuli receiving a mean rating lying in the interval 36.1–67 mm. With one single exception, all tones intended by the singer subject as “twang” were classified as “twang” by the listeners. The standard deviations varied between 6.8 and 27.6 mm. These results support the conclusion that the examples produced by the singer were representative and also that the expert listeners found it reasonably easy to decide if the tones were sung in “twang” or not. Thus, the test offered a quantitative evaluation of how typical the individual examples were of what the expert panel regarded as “twang.”

Spectrum

Figure 4 shows the long-term average spectrum (LTAS) of the two text versions of the song. “Twang” differed from neutral with regard to the overall level. Furthermore, there is a high peak near 1900 Hz in “twang” and a less prominent peak near 1500 Hz in neutral. These peaks appear to correspond to the third and fourth partials, respectively, of the most common F0 in the example, which was about 450 Hz.

Voice source

Figure 5 compares subglottal pressures, SPL and MFDR for identical melody tones sung in the “twang” and neutral styles. In most cases, the pressure was clearly higher in “twang” than in neutral, the average difference amounting to 2.6 cm H2O. As expected as well, SPL and MFDR differed in a similar way—average 6.4 dB and 334 L/s². Qclosed was higher in “twang” (Figure 6A), whereas the AC amplitude (Figure 6B) and H1 − H2 (Figure 6C) tended to be lower. NAQ (Figure 7), being the normalized ratio between AC amplitude and MFDR, was mostly lower in “twang.”

Formant frequencies

The inverse filter analysis is based on information about the formant frequencies. Our analysis concerned the vowel /æ/ only. The formant frequencies of the singer’s performance of Ain’t

| Table 1. Coefficient of Determination ($R^2$), Constant (Slope), and Intercept of a Linear Regression Analysis of the Correlation Between the Six Listeners’ Ratings of “Twanginess” in Repeated Stimuli |
|---|---|---|---|---|---|---|
| Rater | 1 | 2 | 3 | 4 | 5 | 6 |
| $R^2$ | 0.866 | 0.937 | 0.843 | 0.350 | 0.684 | 0.752 |
| Slope | 0.856 | 0.973 | 0.931 | 0.564 | 0.896 | 1.042 |
| Intercept | 19.4 | 6.8 | 0.3 | 16.9 | 3.1 | 6.2 |

FIGURE 3. Rank-ordered mean ratings of “twanginess.” The bars represent one standard deviation. Filled circles and triangles refer to tones intended by the subject as neutral and “twang,” respectively.
nobody using the actual lyrics were also evaluated using the same Decap program. The results are shown in Figure 8. For all vowels, $F_1$ and $F_2$ were higher, whereas $F_3$ and $F_5$ were lower in “twang.”

Multiple regression analysis
The results suggest that our subjects’ examples of “twang” and neutral differed in several aspects of both voice source and formant frequencies. It is then interesting to find out the significance of the various parameters to the mean twanginess ratings.

Table 2 lists the correlations between the mean twanginess ratings and the various parameters analyzed. As can be seen in the table, all parameters except $F_4$ showed a significant correlation. $F_1$, SPL, $H_1 - H_2$, and $Q_{Closed}$ showed the highest correlations, whereas $F_3$, AC amplitude, MFDR, and $F_3$ showed the lowest.

A way of examining the significance of the various parameters in the mean twanginess ratings is to run a multiple regression analysis on the mean ratings and the parameters. Such an analysis was run post hoc by means of the SPSS software (Statistical Package for the Social Sciences). Table 3 shows
the results. According to the analysis, 87% of the variation of the mean ratings could be explained by the parameters analyzed. $Q_{\text{Closed}}$, $H_1 - H_2$, AC amplitude, $F_1$, and $F_2$ were all significantly correlated with the mean twanginess rating. As shown by the standardized regression coefficients Beta in Table 3, the rank order of the importance of the parameters was $F_1$, $H_1 - H_2$, $Q_{\text{Closed}}$, $F_2$, and AC amplitude. The perception of “twang” increased with increasing $Q_{\text{Closed}}$, $F_1$, and $F_2$, and with decreasing $H_1 - H_2$ and AC amplitude.

Figure 9A–C illustrates the relevance of the source and formant parameters in the mean “twanginess” ratings in terms of scatter plots including trendlines. The source parameters alone produced an $R^2$ of 0.788, the formant parameters alone produced 0.911, and the source and formant parameters combined produced 0.945. Thus, the formant parameters were more influential on the mean “twanginess” ratings than the source parameters. It should be recalled, however, that these results concern the vowel /æ/ only.

### DISCUSSION

The listening test suggested that our subject produced examples that were rather easily identified as “twang” and “not twang,” respectively. Hence, they seemed to be reasonably representative. However, the term “twang” may not have exactly the same meaning among different expert groups. For example, according to the Oxford Dictionary, “twang” is a “distinctive nasal pronunciation characteristic of the speech of an individual or region.” However, Estill (2005) distinguishes between nasal and oral “twang.” The “twang” studied in the present investigation reflects the idea of “twang” of our expert listeners.

“Twang” has even been described as a characteristic of classically trained operatic singers (eg, Ref. 1). This particular meaning of the term seems closely related to the presence of a singer’s formant cluster. However, the voice source of operatic singing differs from that observed for “twang” in the present study. For example, the mean NAQ value in classical singing has been found to amount to about 0.16, whereas the NAQ values observed here for the “twang” quality hovered around 0.1. It can also be mentioned that the formant frequency characteristics of “twang” observed here differed radically from those typical of operatic singing. Hence, the type of “twang” considered in the present study is not a characteristic of classically trained operatic singing.

With regard to the voice source, Titze et al. (2003) found that a short open phase contributed significantly to the “twanginess” ratings. In their listening tests, they varied the open quotient in the range 0.3–0.8, that is, the $Q_{\text{Closed}}$ varied between 0.7 and 0.2. In our data, the maximum $Q_{\text{Closed}}$ amounted to 0.513. The reason for the substantially higher values in their study may be that $F_0$ in their syntheses seemed to be about 200 Hz, whereas our female singer was singing at considerably higher pitches, in the $F_0$ range 330–520 Hz. In a female subject, this implies thinner vocal folds, which in turn, should be associated with shorter closed phase. Titze et al. also found that a narrowing and

### TABLE 2.
Correlations Between Mean Twanginess Ratings and the Various Parameters Analyzed

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$r$</th>
<th>Significance (2-Tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MV resp</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$F_1$</td>
<td>0.915</td>
<td>0</td>
</tr>
<tr>
<td>SPL</td>
<td>0.798</td>
<td>0</td>
</tr>
<tr>
<td>$H_1 - H_2$</td>
<td>-0.791</td>
<td>0</td>
</tr>
<tr>
<td>$Q_{\text{Closed}}$</td>
<td>0.779</td>
<td>0</td>
</tr>
<tr>
<td>$F_2$</td>
<td>0.689</td>
<td>0</td>
</tr>
<tr>
<td>NAQ</td>
<td>-0.545</td>
<td>0</td>
</tr>
<tr>
<td>$F_5$</td>
<td>-0.501</td>
<td>0.001</td>
</tr>
<tr>
<td>AC ampl</td>
<td>-0.469</td>
<td>0.002</td>
</tr>
<tr>
<td>MFDR</td>
<td>0.388</td>
<td>0.016</td>
</tr>
<tr>
<td>$F_3$</td>
<td>0.335</td>
<td>0.03</td>
</tr>
<tr>
<td>$F_4$</td>
<td>0.076</td>
<td>0.631</td>
</tr>
</tbody>
</table>

**Abbreviations:** $r$, Pearson correlation; MV resp, mean value of the twanginess ratings; $F_1$, $F_2$, $F_3$, and $F_5$, formant frequencies 1–5; SPL, sound pressure level; $H_1 - H_2$, level difference between first and second partial of the voice source spectrum; $Q_{\text{Closed}}$, closed quotient; NAQ, normalized amplitude quotient; AC ampl, AC amplitude; MFDR, maximum flow declination rate.

### TABLE 3.
Output of the SPSS Multiple Regression Analysis of the Relationships Between the Mean Twanginess Ratings and the Indicated Voice Source Parameters (See Text)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Unstandardized</th>
<th>Standard Error</th>
<th>Beta</th>
<th>Standardized</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-97.7734</td>
<td>32.21893</td>
<td>-3.03466</td>
<td>0.004453</td>
<td></td>
</tr>
<tr>
<td>$Q_{\text{Closed}}$</td>
<td>86.6318</td>
<td>40.80957</td>
<td>2.122831</td>
<td>0.040713</td>
<td></td>
</tr>
<tr>
<td>$H_1 - H_2$</td>
<td>-1.92085</td>
<td>0.424891</td>
<td>-4.52082</td>
<td>0.000064</td>
<td></td>
</tr>
<tr>
<td>AC ampl</td>
<td>-18.4112</td>
<td>9.220523</td>
<td>-1.99676</td>
<td>0.053455</td>
<td></td>
</tr>
<tr>
<td>$F_1$</td>
<td>0.092539</td>
<td>0.016894</td>
<td>5.477529</td>
<td>0.000003</td>
<td></td>
</tr>
<tr>
<td>$F_2$</td>
<td>0.030699</td>
<td>0.013869</td>
<td>2.213531</td>
<td>0.033285</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** $F_1$ and $F_2$, formant frequencies 1 and 2; $H_1 - H_2$, level difference between first and second partial of the voice source spectrum; $Q_{\text{Closed}}$, closed quotient; AC ampl, AC amplitude.

* Dependent variable: mean value of the twanginess ratings.
a shortening of the pharynx promoted the impression of “twang.” A shortening of the pharynx, necessarily accompanying a larynx elevation, and a pharyngeal narrowing will increase the $F_2$. Thus, our finding that a high $F_2$ belonged to the characteristics of the “twang” style is in good agreement with the observations reported by Titze et al. Incidentally, the effect of a shortening of the vocal tract would be further enhanced by a retraction of the corners of the lips, a trait typically seen in many singers using “twang.” Moreover, the observed lowering of $F_3$ suggests that “twang” is produced with a front cavity between the tongue tip and the lower incisors. This is likely to be accompanying a retraction of the tongue, needed for a narrowing of the pharynx.

$Q_{\text{Closed}}$ is affected by several factors, such as subglottal pressure, glottal adduction, and often, also by $F_0$. An increase of glottal adduction implies a change of the phonation type along the breathy–pressed continuum, decreasing the distance to the latter extreme. The higher values of $Q_{\text{Closed}}$ that we observed in “twang” are unlikely to depend on a higher subglottic pressure, because this pressure was not consistently higher in “twang” than in neutral for identical $F_0$, as can be seen in Figure 5. Therefore, the higher values of $Q_{\text{Closed}}$ suggest that our subject produced her “twang” examples with a higher degree of glottal adduction, that is, with a phonation type somewhat closer to the pressed extreme than the neutral examples. The differences between “twang” and neutral with regard to AC amplitude, $H_1 - H_2$ difference, and NAQ all support the same conclusion.

The sound pressure SPL of the tones produced in “twang” were, on average, 6.4 dB higher than those produced in a neutral style of singing. A relevant factor underlying this difference would be the formant frequencies; the higher $F_1$ and $F_2$ values and the lower $F_3$ values observed in “twang” should increase SPL owing to a resonatory effect. Measurements on the Svante Granqvist’s custom-made formant synthesis software MADDE (available at http://www.speech.kth.se/music/downloads/smptool/) revealed that the averaged formant frequencies observed for neutral and “twang” for the vowel /æ/ generated an SPL difference of 5.1 dB between the two. As resonatory effects entail no costs in vocal effort, this seems like a strategy that is advantageous from a vocal hygiene point of view. Interestingly, female and male classically trained western opera singers have developed entirely different resonatory strategies, which also both add to the audibility of the voices in the presence of loud accompaniment sounds.

Our results emanated from one single subject; hence, the generality of the results is an important issue. To find an answer, we ran a test where listeners, all well acquainted with this type of singing, were asked to rate how representative different tones were of “twang.” The results supported the assumption that the “twang” material tested was quite representative of the “twang” style. This is not to claim that a different expert panel would necessarily have produced the same ratings. However, our analyses of subglottal pressure, flow glottograms, and formant frequencies appear to agree with Titze et al.’s results derived from synthesis. It is worthwhile to recall that Titze et al.’s expert panel was American, whereas ours was Swedish. The similarity in our findings suggests that the type of “twang” that we analyzed was similar to that analyzed by Titze.

The multiple regression analysis revealed that $F_1$, $H_1 - H_2$, and $Q_{\text{Closed}}$ were particularly important factors for the perception of twanginess. The general relevance of $F_1$ and the other formants, however, is not reflected in this analysis, because it was limited to the vowel /æ/. Figure 8 shows a more relevant result, in which, almost without exception, $F_1$ and $F_2$ are higher in “twang” than in neutral, whereas $F_3$ and $F_5$ are lower. Thus, “twang” seems to be characterized by both phonatory and resonatory properties.

**CONCLUSIONS**

This single-subject investigation has revealed some vocal characteristics of “twang” quality. Our subject produced typical examples of “twang” and neutral voice qualities that were correctly classified by a panel of experts. Subglottal pressure was often higher, whereas SPL was always higher in “twang” than in neutral. With regard to the voice source, $Q_{\text{Closed}}$ was greater, whereas AC amplitude and $H_1 - H_2$ were smaller, and NAQ tended to be lower in “twang.” With regard to formant frequencies, $F_1$ and $F_2$ tended to be higher and $F_3$ and $F_5$ lower in “twang.” The formant differences were more important than the source difference for the perception of “twanginess.” Moreover, some of the higher SPL values in “twanginess” seemed to be
caused mainly by resonatory effects owing to the higher $F_1$ and $F_2$. As resonatory effects occur independent of the voice source, the formant frequencies in “twang” may reflect a strategy advantageous from the point of view of vocal hygiene.

Acknowledgments

The kind cooperation of our subject is gratefully acknowledged, as is the valuable assistance of Peter Branderud and Hassan Djamshidpey at the Linguistics Department of the Stockholm University during the recordings. Financial support for this Research and Development project was provided by the University College of Music Education in Stockholm (SMI) Sweden.

REFERENCES