Monkey See, Monkey Do?  
The Effect of Nonverbal Conductor Lip Rounding on Visual and Acoustic Measures of Singers’ Lip Postures

James F. Daugherty¹ and Melissa C. Brunkan²

Abstract
If singers, without prior prompting, mimicked a conductor’s nonverbal behavior and if this mimicry changed their vocal sound in less than a second, then such a phenomenon could interest vocal music teachers as a time-efficient pedagogical strategy. We tested this claim (“What they see, you will get”), which appears in choral methods literature, by measuring visual and acoustic responses to one nonverbal conductor behavior in a particular singing context. Specifically, we sought to determine whether singers (N = 114) performing the first phrase of Mozart’s motet, “Ave Verum Corpus,” would mimic a conductor’s rounded lip posture on two /u/ vowels. We also wondered whether conductor lip rounding affected these singers’ tone quality. Visual measures (within-subjects photo comparisons and photo grid analyses) indicated that more than 90% of participants displayed more lip rounding on both /u/ vowels in the experimental condition as compared with baseline. Formant frequency profiles indicated that more than 90% of singers lowered all four examined formant frequencies each time the conductor rounded his lips. We discussed these converging visual and acoustic data in terms of the study’s limitations, potential pedagogical implications of mimicry by vocal performers, and directions for future research.

¹University of Kansas, Lawrence, KS, USA  
²Louisiana State University, Baton Rouge, LA, USA

Corresponding Author:  
James F. Daugherty, Division of Music Education and Music Therapy, University of Kansas, Murphy Hall 448, 1530 Naismith Drive, Lawrence, KS 66045-3103  
Email: jdaugher@ku.edu
Keywords
choral pedagogy, mimicry, nonverbal behavior, lip rounding, singing pedagogy

Some choral pedagogy materials exhibit the author’s testimonial belief that singers will mimic particular nonverbal conductor behaviors. A best-selling instructional video, for example, argues, “The whole body is the conducting gesture” and “What they (the singers) see is what you get” (Eichenberger & Thomas, 1994).

For almost a century, various iterations of this belief have appeared in choral methods books. Gehrkens (1919), for instance, asserted that the “psychological basis of conducting” rests on an assumption that “human beings have an innate tendency to copy” a conductor’s actions, “often without being conscious that they are doing so” (p. 3). Krone (1949) suggested that if a conductor exhibits good posture, students will display the same “through unconscious imitation on their part” and “less nagging” on the conductor’s part (p. 4). Similarly, Jordan (1996) said that choirs “will mirror the posture of their conductors” (p. 12). Bertalot (2002) contended, “If the conductor frowns, they [the singers] frown” (p. 35). To date, however, few controlled studies have tested such beliefs in particular singing contexts.

Several studies have indicated a possible neurobiological basis for imitative behaviors. Di Pellegrino, Fadiga, Fogassi, Gallese, and Rizzolatti (1992) found that neurons in the inferior frontal cortex of a macaque monkey brain fired not only while the monkey was performing grasping activities but also while the animal simply was observing a human being perform grasping actions. Fadiga, Fogassi, Pavesi, and Rizzolatti (1995) found that human participants displayed significantly increased levels of motor evoked potentials when observing particular actions. Various functional magnetic resonance imaging studies showed that observing limb and mouth movements activated frontal cortex motor areas (e.g., Buccino et al., 2001; Iacoboni et al., 1999).

Other investigations have addressed the social dimensions of human mimicry. Chartrand and Bargh (1999), for example, found that participants engaged in a task with a confederate (a research assistant posing as another participant in the study) tended to imitate the confederate’s behaviors. Van Baaren (2003) determined that participants mimicked a confederate to a greater extent in context-dependent tasks than in non-context-dependent tasks and that there was a positive, significant correlation between mimicry in general and scores on an administered test of context dependency. In other experiments, Van Baaren (2003) found that mimicry increased prosocial behaviors even when there was no specific or special relationship between the mimicker and the mimicked.

Authors of various studies examined mimicry of facial and mouth expressions. Meltzoff and Moore (1977) found that infants, some as young as 42 min, spontaneously imitated mouth opening and tongue protrusion. Two studies (Fadiga, Craighero, Buccino, & Rizzolati, 2002; Watkins, Strafella, & Paul, 2003) indicated that participants who listened to or watched speech evidenced increased mouth muscle potentiation. Berger and Hadley (1975) found that observing a person stutter activated lip
muscles. Livingstone, Thompson, and Russo (2009) showed that musicians who watched another musician sing activated expressive facial muscles. Dimberg and Thunberg (1998) found that participants engaged facial muscles within 300 to 400 ms after observing stimulus photos. Dimberg, Thunberg, and Elmehed (2000) concluded that such facial engagement occurred without conscious awareness.

Two studies to date have focused on potential singer mimicry of nonverbal conductor behaviors. Manternach (2011) explored effects of conductor preparatory gestures and conductor head and shoulder movements on the upper-body movement of individual choral singers. Results indicated apparent differences in singers’ head and shoulder movements associated with both the direction of the conductor’s preparatory gesture and the head and shoulder movements displayed.

In a pilot study, Daugherty and Brunkan (2009) explored the effect of rounded conductor lip posture on the lip postures of individual singers (N = 62) during two sung /u/ vowels. Expert ratings indicated that all participants mimicked the conductor’s lip rounding on at least one of the experimental-condition /u/ vowels. However, the researchers viewed these results with caution because the expert panel made decisions while watching performances of full phrases. Some judges reported difficulty remembering the baseline condition by the time they encountered the experimental condition. Moreover, this pilot study did not counterbalance presentation orders.

Choral pedagogy literature traditionally has recommended singing /u/ vowels with “lips pursed and rounded” (Roe, 1994, p. 92), lips “away from the teeth (Hylton, 1995, p. 23), or lips “purse[d] out a bit, relaxed and rounded” (Neuen, 2002, p. 52). Some pedagogues have suggested that a rounded /u/ vowel produces a “richer, and less brilliant color,” the mastery of which “enhances blend” and develops “head voice” (Collins, 1999, p. 309). Others (e.g., Neuen, 2002) have asserted that teaching a rounded /u/ vowel should be among the first steps employed toward improving ensemble sound.

Voice scientists (e.g., Sundberg, 1987; Titze, 2000) have found that changes in vocal tract articulators, such as the lips, potentially alter the dimensions of the vocal tract and hence may modify vocal sound. We hypothesize that such would be the case if singers mimicked conductor lip rounding. Yet that possibility has not been tested with a large group of singers of various ages and singing abilities, who, already articulating /u/ vowels in performance of a sung phrase, might round their lips still more. Vocal music educators would be interested in knowing not only whether singers mimicked conductor-teacher lip rounding without prior instruction to do so but also whether such mimicry demonstrably affected singer sound in a realistic singing task.

**Purpose and Research Questions**

The purpose of this study was to assess visually (by expert ratings of paired photographs and by photo grid analysis) and acoustically (by formant frequency profiles) the lip postures of individual singers (N = 114) as they performed the first phrase melody of Mozart’s “Ave Verum Corpus” while watching a videotaped conductor in
two nonverbal conditions: (a) standard conducting gestures with neutral facial affect (baseline condition) and (b) the same conducting gestures and facial affect with the exception of conductor lip rounding on the /u/ vowels of verum and corpus (experimental condition).

The following research questions guided this investigation: (1) Will singers visually mimic conductor lip rounding in the experimental condition? (2) Will singers’ formant frequency profiles on targeted /u/ vowels vary significantly between baseline and experimental conditions? (3) Will singers recognize that the conductor rounded his lips on /u/ vowels?

**Definitions**

**Formant frequency profile.** Formants are resonance regions of the vocal tract. Movements of the tongue, jaw, velum, and lips, as well as variances in larynx height, modify these resonance regions. A formant frequency profile reports resonance frequency characteristics.

**Mimicry.** Mimicry, as used in this study, is a superficial, yet readily recognizable and therefore measurable, inclination to copy an exhibited stimulus (resemblance behavior) but does not necessarily entail exact replication of the stimulus (replicative behavior).

**Lip rounding.** The term lip rounding, for purposes of this study, will be used in reference to both a circular-like narrowing of the mouth opening and a protrusion of the lips. Although the pedagogical literature does not always distinguish them, there are discrete physiological differences between lip rounding and lip protrusion. As Titze (2000) comments, however, “you probably have to round your lips in order to protrude them because there is only so much tissue to work with” (p. 179). Lip rounding and lip protrusion, whether separately or in concert, share an acoustic effect; both reduce formant frequencies (Sundberg, 1987).

**Method**

**Participants**

**Singers.** Singer participants (N = 114) constituted a convenience array of varied age, sexes, and singing experience, recruited by word of mouth from students, faculty, and staff at a major university. There were 78 (68.42%) female participants and 36 (31.58%) male participants. Participants ranged in age from 14 years (n = 1) to 62 years (n = 2), with a modal age range of 20 to 22 years of age. Singer ages by decade included teens (n = 17, 14.91%), 20s (n = 61, 53.51%), 30s (n = 16, 14.04%), 40s (n = 10, 8.77%), 50s (n = 8, 7.02%), and 60s (n = 2, 1.75%).

Reported years of regular choir singing divided participants into two main groups: (a) 76 (66.67%) had 3 or more years’ experience singing regularly in a choir since early adolescence (middle or junior high school age); (b) 38 participants (33.33%) reported 2 or fewer years’ experience. Among this group with less choral singing
experience, 35 participants (30.70%) had no regular experience singing in any choir since elementary school.

We informed singer participants that the general purpose of the study was “to measure various characteristics of choral singing.” We did not apprise them beforehand of the specific independent variable of interest (lip rounding).

**Expert judges.** A panel ($N = 7$) of expert judges also participated in the study. Each judge regularly had taught private voice lessons, and each judge had directed a choral ensemble. Judges ranged in age from 31 to 54 years. Modal years of judges’ teaching experience was 22 years. Six judges held graduate degrees (doctoral, $n = 3$; master’s, $n = 3$) in voice performance or choral pedagogy. Five judges taught at the university level. Each had taught in public schools.

**Survey Instruments**

Singer participants completed two short surveys. One survey, completed after presentation of a signed institutional research board–approved consent form, solicited demographic information. Participants completed a short exit survey immediately after singing. The survey consisted of one question: “Did you note any differences between what the conductor did in the first and second videos? If not, write none. If so, please briefly describe.”

**Music Excerpt**

The opening phrase of the melody line in Mozart’s motet “Ave Verum Corpus,” sung from memory, constituted the singing task for this study. We chose this particular motet because its opening phrase had two /u/ vowels, the first of which occupies two pitches (thus affording us the opportunity to measure the second pitch /u/ vowel without its being preceded by a consonant) and the second of which is sustained (thus affording us the opportunity to measure at the midpoint of a longer vowel). The two target /u/ vowels, moreover, differed in pitch by only one half step. We used an excerpt sung in Latin because Latin has only five vowels and no diphthongs, and we thought singing in Latin (by reference to either the printed Latin text or provided phonetic spellings) would control for potential variations in /u/ vowel articulation attributable to varying regional English dialects.

**Preliminary Procedures**

Singers heard the excerpt played once on a keyboard while a metronome sounded (MM = 80), as they listened and followed either the Latin or phonetic text. We then asked participants whether they could sing the excerpt a cappella with text as the metronome sounded. We invited those who responded affirmatively to demonstrate that skill. If they did so successfully, we asked whether they could sing the phrase from memory and again invited them to demonstrate.
Participants unsuccessful in singing the excerpt either a cappella or from memory were afforded help in learning those tasks. At no time, however, did we sing, speak, or mouth the text. Rather, we hummed or played the melody on the keyboard until participants reported readiness to sing the excerpt a cappella and from memory and could demonstrate such.

Seven individuals (4 males, 3 females; each of them with less choral singing experience) exhibited erratic pitch behavior, sang the excerpt an octave lower, or both. Each, however, could sing the excerpt from memory with correct rhythms. We retained these persons for the photo portion of this investigation, but we excluded them from the acoustic analyses. Thus, the sample for acoustical analyses consisted of 32 men and 75 women ($N = 107$).

**Stimulus Videotape**

Videotaped conducting served as a control for potential variability in conductor behaviors and to ensure that all participants were responding to the same stimuli. The conductor used a metronome (MM = 80) and a mirror, thereafter examining freeze-framed comparisons within and between conditions, in preparing the videotapes ultimately chosen for the study. See the appendix, Example 1, available in the online supplemental material (http://jrme.sagepub.com/supplemental).

A panel of three experienced choral conductors reviewed the final stimulus tape for consistency of facial affect and conducting gestures. Each evaluator attested that the only change noticeable was the addition of lip rounding in the experimental-condition /u/ vowels.

**Performance Protocol**

During performances, we projected the stimulus videotape so that the conductor appeared life-size, as determined by having him stand beside the projected image prior to the study. Singers stood on a line affixed to the floor to ensure a consistent distance between them and the conductor. Before each condition, participants heard the starting pitch on a pitch pipe. Because choral conductors rarely have reason to seek less singer lip rounding on /u/ vowels, all participants viewed the baseline conducting video first, followed by the experimental condition video in which the conductor displayed rounded lips on the /u/ vowels of *verum* and *corpus*. Video 1 provides an example of this performance protocol with 1 participant (available at http://jrme.sagepub.com/supplemental).

**Preparation of Participant Photos**

A digital RCA Small Wonder EZ2000 camera recorded performances in Audio Video InterLeaved (.avi) format at 29.97 frames per second (fps). We positioned the camera at a consistent distance of 16 in. from participants and at a slight angle to avoid direct
sight line interference with the conductor. The camera was zoomed slightly, if needed, so that we could accommodate varying heights of participants. All other camera settings remained consistent for each participant.

Video recordings were transferred digitally to a MacBook computer for viewing with QuickTime software. We used the frame number function of the QuickTime Player to identify the target /u/ vowels (verum second pitch and final corpus) in both baseline and experimental conditions. The duration of each targeted verum /u/ was 0.47 s; the duration of the sustained corpus /u/ was 1.30 s. Thus, at 29.97 fps, the duration of the verum /u/ was approximately 14 frames, and the duration of the corpus /u/ was approximately 39 frames. Each frame constituted approximately 0.033 of a second, or 3.3 centiseconds (cs). We made a still-picture screenshot at or near the midpoint of each of the target /u/ vowels. In some cases, we allowed a leeway of ±5 frames (up to 16.5 cs) from midpoint because (a) we desired similar eyelid positions in each pair of photos (open eyes in one photo and semiclosed eyes in the other conceivably could siphon judges’ attention from the lips) and (b) our research question was whether lip postures would change during experimental-condition /u/ vowels, not whether lip postures would change at precisely the same moment.

We incorporated the photos into a Microsoft PowerPoint slide show for viewing by the expert panel. Each slide contained four photos in two rows: (a) The upper row included full head baseline and experimental-condition target vowel photos, and (b) the lower row included cropped photos (taken from the same full head photos appearing on the same slide) of participants’ lips only in each condition. Because there were 114 singers with two target vowels in two conditions, each participant was featured in two slides (verum conditions and corpus conditions). We used only the crop function to feature participants’ faces in the first row of PowerPoint photos on each slide and to obtain the second row, lip-only photos; the resize drag function was never used.

Before presentation to the judges, we randomly ordered these 228 slides using a random numbers table. Thereafter, again using a random numbers table, we reversed presentation order for one half of the slides (n = 114), with the experimental-condition photo pairs appearing first and the baseline photo pairs appearing second.

**Expert Ratings: Visual Analog Scale**

Judges separately viewed the PowerPoint presentation in slide show mode on a 27-in. Dell 3007WFP-HC Flat Panel Monitor (2,560 by 1,600 pixels, 32-bit color) connected to a Dell Optiplex 745 computer. The first photos (Condition 1 head shot with lip shot below it) on each slide appeared immediately as each judge transitioned to a new slide. Thereafter, the second photos (Condition 2 head shot with lip shot below it) appeared after an interval of 1.5 s.

Judges were directed to view the first (left) photo in each pair of photos as a baseline, focusing on each participant’s lip rounding, and then to evaluate the lip posture in the second (right) photo by marking a visual analog scale (VAS). The VAS had equidistant line lengths on each side of a midpoint, N.C. (no change) tick. Judges
circled N.C. if they discerned no change in the second photo. Otherwise, they made a vertical line to indicate how much less or how much more lip rounding they observed in the second photos. Members of the expert panel could view each complete slide for as long as desired to make an evaluation and thus controlled when they wished to transition to a new slide.

We measured markings in millimeters from the N.C. line and recorded them on a spreadsheet for subsequent analyses. For stringency, we decided a priori not to count as evidence (no change, less lip rounding, more lip rounding) photo pairs that elicited less than unanimous expert agreement.

**Grid Analyses**

**Procedure.** As a test of experts’ perceptual ratings, we used Dr. Levin’s Phi Dental Grid software (Levin & Meisner, 2010) for photo grid analysis. Developed for use in dental aesthetics, this software featured a grid overlay option amenable to user specifications. We superimposed a grid of 4-mm squares over a sample of the same photo pairs viewed by the expert panel, using judges’ PowerPoint slides in normal presentation mode at 150% zoom across all pairs. We chose this protocol because toggling between grid and photo was problematic in full PowerPoint slide show view mode, and we found that squares of 4 mm permitted easiest reading and calculation.

We used two grid measurements for intraindividual comparison: (a) a horizontal measurement, counting the number of grid squares (or quartile divisions thereof) from the right edge corner of a participant’s lips to the left edge corner of the lips, and (b) a vertical measurement, counting the number of grid squares (or quartile divisions thereof) from the center bottommost outer edge of a participant’s lips to the center topmost outer edge of the lips.

We positioned the grid over each photograph, then moved it as needed so that a grid line aligned with the right lip corner (for horizontal measurements) or the center outermost edge of the bottom lip (for vertical measurements). For lip postures that exceeded the last complete grid either horizontally or vertically, we applied a centimeters-millimeters ruler to the computer screen to measure any remaining distances of 1 to 3 mm (see Example 2 available at http://jrme.sagepub.com/supplemental).

**Sampling.** We employed two sampling procedures for grid analysis. First, to analyze approximately one fourth \( n = 56, 24.56\% \) of all photo pairs, we selected by means of a random numbers table 28 verum photo pairs and 28 corpus photo pairs with no duplication of participants. This stratified random sampling procedure yielded analysis of almost half of all participants \( n = 56, 49.12\% \) in one of their vowel performances. Among this sample were 37 females (47.44% of all female participants) and 19 males (52.78% of all male participants) as well as 21 less chorally experienced singers (55.26% of all participants in that category) and 35 more chorally experienced singers (46.05% of all participants in that category).

Second, we decided to employ purposive sampling, as needed, to ensure that all photo pairs with less than unanimous agreement from the expert panel would be
included in the sample subject to grid analysis. Thus, any such pairs not among those 56 pairs selected by stratified random sampling subsequently were added.

To enhance the credibility of results, we decided (a) to count as evidence of “more lip rounding” only those photo pairs where both horizontal and vertical grid measurements in the experimental condition exceeded those of the baseline condition and (b) to use an independent reliability assessor. Thus, one of the researchers and an independent assessor performed the same grid analysis protocol on all selected photo pairs. Obtained reliabilities (agreements divided by agreements + disagreements) were .98 (horizontal measures) and .97 (vertical measures).

**Acoustic Analysis**

We used a digital Edirol R-09 audio recorder in .wav format at a sampling rate of 44.1 kHz and consistent volume setting to obtain data for acoustic analyses. We positioned the recorder at a consistent distance of 6 in. from the left corner of participants’ mouths to avoid the direct air stream and to prevent masking participants’ lips in the video recordings.

After isolating the sung /u/ vowels of *verum* and *corpus* in both baseline and experimental conditions, we used the midpoint of the time interval of each vowel to obtain fundamental frequency (F0) and to compile formant frequency profiles using Praat software (Boersma & Weenink, 2010). For formant extraction, Praat applied a Guassian-like window to compute linear predictive coefficients through the Burg algorithm integrated in the software. We used these computations to obtain frequency readings of the first four formants in each of the four /u/ vowels performed by each participant.

**Results**

Results are presented according to the three research questions posed for this study. We used an alpha level of .05 to determine significance on statistical tests.

**Research Question 1: Visual Measures**

Experts examined 228 paired baseline and experimental-condition photos (two different vowel iterations per participant) to determine whether they perceived a change in the second photo of the pair. Ratings indicated significantly more instances of observed lip rounding in the experimental condition than for the baseline, \( \chi^2(1, N = 228) = 109.14, p < .0001 \). By unanimous rater agreement, 103 participants (90.35%) evidenced more instances of lip rounding on both experimental-condition /u/ vowels. Every participant (100.00%) received a unanimous verdict of change (more lip rounding) on at least one of the experimental-condition /u/ vowels.

Mean panel ratings of the magnitude of lip rounding, as measured by the VAS markings, suggested more lip rounding in all 228 experimental-condition photos, but
in five pairs (2.21%), one of the judges dissented; in four pairs (1.75%), two judges dissented; in one pair (0.44%), three judges dissented; and in one pair (0.44%), four judges dissented. Overall mean ratings indicated no perceived change toward less lip rounding.

All participants (n = 38, 33.33%) with less choral singing experience showed, by unanimous panel agreement, more lip rounding on both experimental-condition /u/ vowels as compared with baseline, including the 7 participants excluded from acoustic analyses because of erratic pitch behaviors.

Figure 1 presents judges’ mean ratings (in VAS millimeters) of all 228 experimental-condition photos disaggregated by vowel. A Wilcoxon matched-pairs test indicated no significant difference between ratings for the experimental-condition verum (M = 13.45, SD = 8.40) and corpus (M = 12.46, SD = 7.92) /u/ vowels, z = −.070, p = .48 (two-tailed test).

**Illustrative disaggregation by category of perceived change.** Example 4 (online at http://jrme.sagepub.com/supplemental) shows paired baseline and experimental-condition photos of 3 participants singing the corpus /u/ vowel: (a) 1 participant whom judges rated as evidencing very much change (45 to 49 mm), (b) 1 participant rated as displaying moderate change (20 to 24 mm), and (c) 1 participant rated as showing a little change (1 to 4 mm).

**Grid Analyses**

VAS ratings yielded 11 (4.82%) photo pairs with less than unanimous expert agreement. Four of these pairs already were included among the 56 pairs we selected a
priori to analyze as part of our stratified random sample. We added the remaining 7 pairs with less than unanimous expert agreement to the sample, for a total of 63 photo pairs representing 63 participants (55.26%).

Results indicated that 59 (93.65%) of the 63 participants sampled evidenced more lip rounding in the experimental condition as compared with baseline according to both horizontal and vertical dimensions. All participants in the sample evidenced less distance between lip corners (horizontal grid measurement). “No-change” vertical measurement results \( (n = 4) \) were divided evenly according to vowel \( (\text{verum, } n = 2; \text{corpus, } n = 2) \).

Overall mean differences (in millimeters) between baseline and experimental-condition photos were as follows: (a) horizontal dimension, \( M = 2.22, SD = 1.54 \); (b) vertical dimension, \( M = 1.86, SD = 1.14 \); and (c) combined horizontal and vertical dimensions, \( M = 4.11, SD = 2.25 \). Of the 11 photo pairs where expert scale ratings indicated less than unanimous agreement, 10 (90.91%) of these experimental-condition photos showed both decreased horizontal distance (1 to 4 mm) and increased vertical distance (1 to 2 mm) lip dimensions.

**Research Question 2: Formant Frequency Profiles**

Tables 1 and 2 display results of formant frequency analyses. Men and women differ in average vocal tract length, which affects vocal tract–dependent formant frequencies. Therefore, results are presented according to participant sex.

### Table 1. Male \( (n = 32) \) Fundamental Frequency \( (F0) \) and Formant Frequency \( (F1–F4) \) Profiles by Word and Vowel

<table>
<thead>
<tr>
<th>Word/Vowel</th>
<th>Base ( M ) (SD)</th>
<th>Exp ( M ) (SD)</th>
<th>% Diff</th>
<th>Contra ( N )</th>
<th>( t (df) )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verum</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F0</td>
<td>194.19 (4.51 Hz)</td>
<td>194.85 (± 4.71 Hz)</td>
<td>0.34</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>F1</td>
<td>388.38 (25.23 Hz)</td>
<td>364.04 (32.84 Hz)</td>
<td>6.27</td>
<td>0.00</td>
<td>5.30 (31)</td>
<td>.0001*</td>
</tr>
<tr>
<td>F2</td>
<td>926.33 (88.50 Hz)</td>
<td>825.54 (69.96 Hz)</td>
<td>10.89</td>
<td>0.00</td>
<td>8.74 (31)</td>
<td>.0001*</td>
</tr>
<tr>
<td>F3</td>
<td>2376.59 (148.89 Hz)</td>
<td>2238.23 (110.88 Hz)</td>
<td>5.83</td>
<td>0.00</td>
<td>4.74 (31)</td>
<td>.0001*</td>
</tr>
<tr>
<td>F4</td>
<td>3448.37 (291.62 Hz)</td>
<td>3260.79 (211.03 Hz)</td>
<td>5.44</td>
<td>0.00</td>
<td>3.55 (31)</td>
<td>.0012*</td>
</tr>
<tr>
<td><strong>Corpus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F0</td>
<td>186.97 (5.85 Hz)</td>
<td>187.12 (4.32 Hz)</td>
<td>0.09</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>F1</td>
<td>376.16 (27.29 Hz)</td>
<td>351.38 (26.44 Hz)</td>
<td>6.59</td>
<td>0.00</td>
<td>6.38 (31)</td>
<td>.0001*</td>
</tr>
<tr>
<td>F2</td>
<td>928.79 (80.91 Hz)</td>
<td>839.41 (67.97 Hz)</td>
<td>9.63</td>
<td>0.00</td>
<td>5.97 (31)</td>
<td>.0001*</td>
</tr>
<tr>
<td>F3</td>
<td>2357.95 (95.43 Hz)</td>
<td>2282.64 (107.22 Hz)</td>
<td>3.20</td>
<td>0.00</td>
<td>5.46 (31)</td>
<td>.0001*</td>
</tr>
<tr>
<td>F4</td>
<td>3378.47 (218.83 Hz)</td>
<td>3283.21 (177.21 Hz)</td>
<td>2.82</td>
<td>0.00</td>
<td>3.74 (31)</td>
<td>.0007*</td>
</tr>
</tbody>
</table>

Note: % Diff = mean percentages of change between baseline and experimental conditions; Contra \( N \) = number of participants with contrary change between conditions.

*Significantly lower \( M \) formant frequency in experimental condition, \( p < .05 \), Bonferroni correction = 0.03125.
Within-subjects $t$ tests for correlated samples (two tailed) showed significantly lowered mean frequencies for examined formants in the experimental condition for both verum and corpus /u/ vowels with both sexes, with the sole exception of the female corpus F4 frequency. Of 107 participants examined acoustically, 97 (90.65%) showed lowering of all four formant frequencies on both experimental-condition /u/ vowels. All (100%) male participants evidenced lower frequencies across all four formants in both experimental-condition /u/ vowels as compared to baseline. For both sexes, the largest mean percentages of change occurred with F2.

**Summary of Research Questions 1 and 2 Results: Converging Visual and Acoustic Data**

Formant frequency profile results demonstrated that “something” happened with a significant majority (90.65%) of singers during both experimental-condition /u/ vowel performances to lower frequencies on all four formants examined. Yet we cannot presume by these data alone that changed lip postures played a primary role. Participants, for example, may have lowered larynges with little change in lip posture as they observed the conductor’s lips.

However, when we consider acoustical results in conjunction with visual analyses, which showed more lip rounding on both experimental-condition /u/ vowels by a similarly significant majority of participants (90.35% according to perceptual ratings, 93.65% of the sample for grid analysis), the most parsimonious explanation is that lip rounding likely played the major role. Thus, not only did a significant majority of
participants visually mimic the conductor’s lip posture, but by doing so, they also changed the timbre of their vocal sound.

**Research Question 3: Participants’ Awareness of Conductor Behavior**

Immediately after their performances, participants responded in writing to the prompt regarding differences that may have been noted between the two conductor videos. Response rate was 100% ($N = 114$).

We sorted responses into two sets of exhaustive and mutually exclusive categories. The first set of categories organized all responses according to whether participants noticed a change. The second set of categories arranged responses by accuracy of any changes noted.

*Change/no change.* Ninety-nine participants (86.84%) reported they noticed a change in the conductor’s behavior. Fifteen participants (13.16%) wrote “no change” or “none.”

*Accuracy of noted change.* Seventeen participants (14.91%) noticed an incorrect change in conductor behavior. These noted changes uniformly addressed tempo (e.g., “The conductor had a faster tempo in the second video,” $n = 15$ responses; “The conductor held out the last note longer in the second video,” $n = 2$ responses). These responses were categorized as incorrect because the conductor used a metronome through the final cutoff point in both conditions.

About half of all participants ($n = 56, 49.12$%) reported a change in generally correct but nonspecific terms. These responses uniformly described a change related to the conductor’s mouth (e.g., “He mouthed the words in the second video,” “He helped me remember the words the second time”) yet did not specifically address conductor behavior with respect to /u/ vowels.

Twenty-six (22.81%) participants reported accurately and specifically that the conductor did something with his lips exclusively on experimental-condition /u/ vowels (e.g., “He rounded his lips on the oo vowels,” “He focused on the oo vowel anytime I said ‘rum’ or ‘pus’”). Most participant responses ($n = 88, 77.19$%), then, were not both accurate and specific.

**Discussion**

The major findings of this study are that (a) 103 (90.35%) of the 114 singer participants twice mimicked a conductor’s rounded lip posture according to unanimous expert visual assessment (as confirmed by photo grid analysis of two dimensions of lip posture in a selected sample of participants, 93.65%) and (b) 97 (90.65%) of 107 participants twice exhibited lowering of all four examined formant frequencies when the conductor rounded his lips. These results raise matters that merit further research and reflection by vocal music educators.

The question, “So what?” comes immediately to mind. In that regard, it may be helpful to remember that for singers, changes in positioning of the vocal tract articulators alter the shape and acoustic properties of the vocal tract and hence change the
perceived timbre of the vocal sound. In this study, mimicry of conductor lip rounding produced significant reduction in mean formant frequencies. These lowered formant frequencies would likely be perceived as a somewhat “darker,” more “covered,” perhaps “richer” sound.

This somewhat darker, more covered sound, depending on aesthetic and other dispositions, might be preferred on sung /u/ vowels. If such is the case, then results from this study may suggest that demonstrating desired lip posture nonverbally could be a time-efficient pedagogical strategy. The conductor in this study, for example, rounded his lips for 0.47 s and 1.30 s, respectively, on each /u/ vowel. Singers responded with increased lip rounding, thus changing their vocal sound, during those very brief intervals of time.

We find such data intriguing. But more research is needed for one to refute or confirm these findings. Although participants represented an array of ages and previous singing experiences, and although it may not be feasible to gather a truly random sample of singers because singing is a widespread human behavior, results from this convenience population should not be generalized.

This investigation, moreover, isolates one conductor behavior, that is, lip rounding. In everyday contexts, a conductor may exhibit multiple nonverbal behaviors at any given moment. Authors of future studies might test mimicry when conductor lip rounding occurs in concert with other physiological behaviors, such as raised eyebrows, varied facial expressions, or changes in conducting plane and hand shape.

The primary task for singers in this study was to watch the conductor while singing in an arguably “neutral” language (Latin). Study protocols (e.g., memorization of the sung excerpt, standing at the same distance in front of the conductor) assisted fulfillment of that task. In a rehearsal or studio context, however, singers may multitask, stand or sit at various angles from the conductor-teacher, focus their eyes on a score or other environmental landmarks, and sing in their native tongue. Authors of subsequent research might explore singer mimicry in varying situations where attention may not be focused consistently or solely on the teacher. Moreover, by replicating this study with small groups of singers instead of solo singers, researchers could test whether there may be a social aspect to singer mimicry, as Van Baaren (2003) found in other mimicry contexts.

There could be reasonable disagreement about the stringency applied for collecting and interpreting data. It might be argued, for example, that a majority of panel ratings (rather than unanimous agreement) and horizontal-only grid analysis (rather than both horizontal and vertical analyses) would constitute sufficient visual indication of lip rounding or that lip rounding on one vowel (rather than two) suffices for mimicry. Similarly, it could be argued that demonstrated reduction in F4 frequencies for female singers is too strict, given less mobility of that formant in females. To avoid possibly numerous qualifications and to enhance credibility of results in this particular exploration of an underinvestigated phenomenon, we consistently opted for a conservative approach. Future studies might incorporate different research decisions.
Authors of future investigations also might consider other measures for visual assessment. Livingstone et al. (2009), for example, employed electromyography (EMG) and motion capture measures with a small group (N = 11) of singers. Apart from feasibility considerations with the large number of singers in this study, we thought that sensors attached around participants’ mouths would clue singers to the study’s purpose. In future studies, however, researchers might sample discrete lip muscle activity in a manageable number of singers by including distractor sensors placed on participants’ shoulders and chests. In time, moreover, facial recognition software may become sufficiently robust to visually assess singers’ lip movements.

We used expert perceptual ratings in this study because teachers in a typical choir rehearsal or voice studio context make pedagogical decisions based on their perceptions of student behaviors. Yet data from other measures, such as EMG, could be instructive for voice teacher education. We find it interesting, for example, that grid analysis of 11 photo pairs with less than a unanimous panel rating indicates that 10 of those experimental-condition lip postures meet our criteria for rounding. Exploring possible correlations between teacher perceptions and both the degree and kind of lip muscle activity evidenced by particular singers merits research. It may be that some lip anatomies are more difficult for teachers to assess visually than others.

Whereas 82 (71.93%) participants reported general awareness that the conductor did something with his mouth, only 26 (22.81%) of them described the specific dimensions of that changed behavior. In retrospect, we wish we had asked as well a second question, “Were you aware that your own lip posture and vocal sound changed when the conductor rounded his lips?”

To the extent that it can be proved scientifically through ongoing testing and refinement, the overall theory of “what they see is what you get” offers interesting possibilities for vocal pedagogy as an additional strategy for teaching and learning. It also could prompt some changes in the training of vocal teacher-conductors by underscoring a need for increased attention to both (a) how human voices work physiologically and acoustically and (b) how particular teacher-conductor gestures and postures might nonverbally facilitate desired sound changes without compromising the ease and efficiency of vocal production. Not all mimicry in choral-vocal contexts may be beneficial or even benign. Vocal music educators would need sufficiently in-depth physiological understanding of the voice to tell the difference.

Finally, the holistic, interactive perspective inherent to the theory of “what they see is what you get” in vocal music contexts, that is, “the whole body is the conducting (and thus teaching) gesture,” possibly points to phenomena that may not be understood fully with reductionist, scientific methods alone. That factor may complicate the course of empirical research. But it need not detain it. Even simple experiments, such as the present study, can contribute to ongoing dialogue in this arena by assessing the credibility of aspects of the overall theory at particular junctures in particular contexts.

Empirical studies of human mimicry appear in the research literature of other disciplines, such as psychology (e.g., Dimberg et al., 2000), neurobiology (e.g., Fadiga...
et al., 2002), and sociology (e.g., Van Baaren, 2003). Belief that singers mimic certain nonverbal conductor behaviors, particularly, posture, informs some choral methods materials (e.g., Eichenberger & Thomas, 1994; Gehrken, 1919; Jordan, 1996; Krone, 1949). However, to the best of our knowledge, in no previous controlled study have researchers examined from a pedagogical perspective whether singers engaged in vocal performance may mimic a particular conductor facial behavior or whether such mimicry, when present, matters in terms of its acoustical consequences.

Both mimicry and vocal performance are highly complex human behaviors. A line of investigation that involves consideration of these behaviors simultaneously must proceed carefully in a one-step-at-a-time fashion. The present study offers replicable procedures that may be useful to other researchers, and its data contribute some empirical evidence pertinent to an underinvestigated matter of professional discourse and interest. Further investigation in a variety of realistic singing contexts appears warranted.

Acknowledgments
The authors gratefully acknowledge as an inspiration for this study the pioneering reflections on nonverbal choral pedagogy behaviors, articulated during the course of a distinguished career, by Rodney Eichenberger, professor emeritus at Florida State University. We thank Jeremy Manternach, University of Arizona, for preparing the stimulus videos.

Declaration of Conflicting Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

Note
1. The obicularis oris muscle complex, embedded in human lips and encircling the mouth, functions much like a drawstring when it operates to bring the corners of the lips closer together. As a consequence, this muscle action can produce increased center distance between lower and upper lips and/or some protrusion of the mid lips. However, the peripheral and marginal fibers of the obicularis oris complex function somewhat differently in rounding versus protrusion (Honda, Kurita, Kakita, & Maeda, 1995), and the mentalis depressor and labbi oris muscles also come into play with lip protrusion (Murano, Stone, & Honda, 2005). Broadly, lip protrusion entails increased distance between the teeth and the inner surfaces of one or both lips, whereas lip rounding per se need not always do so.

References


**Bios**

**James F. Daugherty** is associate professor of music education and director of the Vocal/Choral Pedagogy research group at the University of Kansas. His primary research interests are lifespan voice pedagogy, choir acoustics, and philosophical foundations of music and music education.

**Melissa C. Brunkan** is assistant professor of music education at Louisiana State University. Her primary research interests include singer gestures, adolescent female voices, and voice pedagogy.