Psychology Faculty Publications | Psychology

2009

The Sensitivity of Auditory-Motor Representations to Subtle Changes in Auditory Feedback While Singing

Dwayne Nicholas Keough
Wilfrid Laurier University

Jeffery A. Jones
Wilfrid Laurier University, jjones@wlu.ca

Recommended Citation

The sensitivity of auditory-motor representations to subtle changes in auditory feedback while singing

Dwayne Keough
Department of Psychology, Wilfrid Laurier University, Waterloo, Ontario N2L 3C5, Canada

Jeffery A. Jones[a]
Department of Psychology and Centre for Cognitive Neuroscience, Wilfrid Laurier University, Waterloo, Ontario N2L 3C5, Canada

Singing requires accurate control of the fundamental frequency (F0) of the voice. This study examined trained singers' and untrained singers' (nonsingers') sensitivity to subtle manipulations in auditory feedback and the subsequent effect on the mapping between F0 feedback and vocal control. Participants produced the consonant-vowel /ta/ while receiving auditory feedback that was shifted up and down in frequency. Results showed that singers and nonsingers compensated to a similar degree when presented with frequency-altered feedback (FAF); however, singers' F0 values were consistently closer to the intended pitch target. Moreover, singers initiated their compensatory responses when auditory feedback was shifted up or down 6 cents or more, compared to nonsingers who began compensating when feedback was shifted up 26 cents and down 22 cents. Additionally, examination of the first 50 ms of vocalization indicated that participants commenced subsequent vocal utterances, during FAF, near the F0 value on previous shift trials. Interestingly, nonsingers commenced F0 productions below the pitch target and increased their F0 until they matched the note. Thus, singers and nonsingers rely on an internal model to regulate voice F0, but singers' models appear to be more sensitive in response to subtle discrepancies in auditory feedback. © *2009 Acoustical Society of America.* [DOI: 10.1121/1.3158600]

I. INTRODUCTION

The role of auditory feedback during vocalization has been a topic of much scientific inquiry. Both speech and singing are debatably the most complex motor actions humans are capable of producing. In order to produce a word, or to sing a musical note, one must possess strict control over respiratory muscles in addition to control over intrinsic and extrinsic laryngeal muscles. This control is achieved by an intricate network of cortical and brainstem areas dependent upon auditory (Sapir *et al.*, 1983; Larson *et al.*, 2008) and proprioceptive (Kirchner and Wyke, 1965; Wyke, 1974; Yoshida *et al.*, 1989) reflex mechanisms. However, nonreflexive systems that utilize auditory feedback also contribute greatly to the development of speech in children (e.g., Oller and Eilers, 1988).

Postlingually, auditory feedback also appears to be monitored and used during ongoing speech. For instance, delaying auditory feedback results in disruptions in the quality of vocal productions (Yates, 1963). Moreover, altering auditory feedback typically elicits compensatory responses in speakers' ongoing vocal productions. Increases in masking noise and decreases in side-tone amplitude (Bauer *et al.*, 2006; Lane and Tranel, 1971), shifts in formant frequencies (Houde and Jordan, 1998; Purcell and Munhall, 2006), and the fundamental frequency (F0) (Burnett *et al.*, 1997; Elman, 1981; Jones and Munhall, 2000; Kawahara, 1998) all cause vocal responses that oppose the respective manipulations. The purpose of the present study is to further our understanding of the role of auditory feedback for the control of F0.

Frequency-altered feedback (FAF) is one paradigm that has been used to examine the importance of auditory feedback for the regulation of voice F0. In a typical FAF study, participants receive auditory feedback regarding their pitch that is higher or lower than their actual vocal productions. Responses to upward or downward perturbations generally result in decreases or increases in F0, respectively. Interestingly, the majority of FAF studies (e.g., Burnett *et al.*, 1997; Burnett *et al.*, 1998; Burnett and Larson, 2002; Elman, 1981; Jones and Munhall, 2000, 2002, 2005; Natke *et al.*, 2003; Toyomura *et al.*, 2007) had participants produce a relative target, but not match a specific frequency, as one must do while singing.

In those studies, most participants compensated for the discrepancy detected between perception and production, although not perfectly (e.g., Burnett *et al.*, 1997; Donath *et al.*, 2002; Larson, 1998; Larson *et al.*, 2000; Natke *et al.*, 2003). On average, manipulations varying from ±100 to 600 cents had a response magnitude of approximately 50 cents (regardless of the direction of manipulation) but ranged from 15–65 cents (cent is a logarithmic unit used to measure small intervals between different frequencies, where 100 cents is equal to 1 semitone). Liu and Larson (2007) observed complete compensation for small perturbations (10 cents), suggesting that the role of the auditory-vocal system is to correct for

―――――――――――――――――
[a]Author to whom correspondence should be addressed. Electronic mail: jjones@wlu.ca

small errors rather than larger F0 deviations. However, it is worth noting that the response magnitudes obtained from previous studies (e.g., Burnett *et al.*, 1997; Donath *et al.*, 2002; Larson, 1998; Larson *et al.*, 2000; Liu and Larson, 2007) were measured by the random presentation of auditory perturbations, whereas the current study assessed sensorimotor adaptation through the presentation of predictable changes in voice F0.

Singing offers researchers a unique window into the study of F0 control. Singing involves producing a succession of musical sounds at a particular (absolute) frequency (e.g., 440 Hz, A4). Thus, singers must maintain accurate vocal control over their F0 to hit the desired notes. Sundberg (1987) (p. 177) reported that if the actual vocal production deviates from the intended target, trained singers are able to compensate and match the tone (A4, 440 Hz) with an accuracy of less than 1 Hz. However, few studies have directly examined the role of auditory feedback while singing (Burnett *et al.*, 1997; Jones and Keough, 2008; Natke *et al.*, 2003; Zarate and Zatorre, 2005, 2008). The results obtained during singing studies that have used FAF (e.g., Burnett *et al.*, 1997; Natke *et al.*, 2003) indicate that the compensation is comparable to those achieved during speech production. The compensatory responses observed in trials of altered feedback indicate that voice production is regulated in a closed-loop fashion (Fairbanks, 1954; Hain *et al.*, 2000; Larson *et al.*, 2008; Lee, 1950). However, laryngeal structures and vocal fold stiffness are set in place prior to vocal onset (e.g., Watts *et al.*, 2003), suggesting a role for open-loop motor planning.

In the case of singing, when learning to produce a target note, stronger reliance on auditory feedback during the initial acquisition stages may be required in order to establish sensorimotor representations that will guide future vocal productions. This idea is based on the premise that the feedback one receives while learning to perform a specific task creates an integrated sensorimotor representation that is directly related to the task (Proteau *et al.*, 1987; Proteau *et al.*, 1992). For instance, Finney and Palmer (2003) demonstrated that the availability of auditory feedback while pianists learned a piece of music significantly improved their ability to play the piece from memory during later recall. However, when pianists were asked to perform well-rehearsed musical sequences from memory, the removal of auditory feedback did not affect musical performance (Finney and Palmer, 2003).

Thus, over time, it is hypothesized that an internal representation is formed or "internal model" that stores the relationships between the motor commands, environment, and sensory feedback for their production (Proteau *et al.*, 1987; Proteau *et al.*, 1992). As a result, pianists may be capable of performing a well-rehearsed piece from memory without auditory feedback because they are relying on internal motor commands that correspond to the musical composition. The same reliance on an internal motor plan may also exist for singing a well-rehearsed piece. For instance, it is possible that vocal productions are initiated based on an increased reliance on the motor plan established for vocalization.

Jones and Keough (2008) investigated whether trained singers, given their extensive training, rely more on a well-established internal representation than nonsingers while

singing. Indeed, identifying differences in F0 control while participants receive subtle manipulations in auditory feedback should elucidate the sensitivity of the underlying mechanisms regulating F0 control. Jones and Keough (2008) found that nonsingers initially compensated to a greater degree than singers when exposed to FAF (−100 cents). However, after this brief exposure to FAF, singers' F0 values were higher than their base line F0 values when they heard their feedback returned to normal. In contrast, no differences were found between the F0 values during the base line and test trials for nonsingers. These effects persisted when singers were required to produce a note other than the one they sang during the altered feedback trials. The results imply that singers rely more on an internal model to regulate their F0 productions during singing than nonsingers.

The present study was designed to investigate the sensitivity of singers' and nonsingers' internal representations to subthreshold (2 cent increments) manipulations in auditory feedback. The data obtained by Jones and Keough (2008) suggested that trained singers compensated less for larger discrepancies between perception and production than nonsingers and instead relied more on their internal models. Nonsingers compensated immediately after exposure to altered feedback of 100 cents, whereas singers required several trials to modify their F0 (Jones and Keough, 2008). Moreover, Zarate and Zatorre (2005, 2008) found that singers were able to successfully ignore the altered auditory feedback (by presumably relying on an internal model) they received and to continue to produce the targets at the desired frequency when instructed to do so. Conversely, nonsingers' vocal productions suggested that they were unable to ignore the altered feedback and as a result adjusted their F0 to compensate for the manipulation (Zarate and Zatorre, 2005, 2008).

Jones and Keough (2008) found that singers initially responded less to large (100 cent) shifts in auditory feedback than nonsingers, who exhibited near perfect levels of compensation almost immediately. However, it remains unknown whether similar F0 values for singers and nonsingers would be obtained using small manipulations in auditory feedback (i.e., 2 cent increments up to 1 semitone, 100 cents). Given the paucity of data, the authors hypothesized that singers and nonsingers would compensate to a similar degree to the altered feedback. However, they expected that singers, due to their extensive training and experience, would compensate for the discrepancy in perception and production more efficiently than nonsingers. That is, the mechanisms that regulate F0 control in singers would be more sensitive to subtle changes in FAF. As a result, singers will reproduce the target notes more accurately than nonsingers while receiving FAF.

Of particular interest was determining when singers' and nonsingers' compensatory responses to small manipulations in auditory feedback would occur. Determining when both groups initiate compensatory responses may also provide an index of how sensitive the underlying mechanisms of the internal representations are to perturbations. Recently, the just-noticeable difference of a single fundamental frequency (120 Hz) to natural sound stimuli (e.g., lowest tone on a violin pitch shifted to 120 Hz; German vowels /i:/ and /a:/

and Italian vowels /i/ and /a/) was shown to be between 4 and 98 cents (Pape and Mooshammer, 2006). Furthermore, Loui *et al.* (2008) found that participants without speech or hearing disorders, and with no formal music training, had a perceptual and production threshold (approximately 2.0 and 2.5 Hz, respectively) that was significantly better than the threshold for amusical ("tone-deaf") participants (approximately 36.2 and 12.3 Hz, respectively). The authors hypothesized that singers would adapt to altered feedback earlier than nonsingers due to their extensive vocal experience achieving pitch targets.

Previous studies using the FAF paradigm have found evidence for sensorimotor adaptation by measuring the mean F0 after feedback was returned to normal (Jones and Munhall, 2000, 2002, 2005; Jones and Keough, 2008). These aftereffects can be observed following exposure to a single trial of FAF (Natke *et al.*, 2003) in nonsingers. To track sensitivity to FAF in this experiment, the authors assessed sensorimotor adaptation by measuring F0 at vocal onset while exposing participants to predictable changes in auditory feedback (increases or decreases in frequency by 2 cents on each successive trial). Using this measure, aftereffects will be evident when voice F0 values start at or near F0 values observed at the end of the previous utterance.

If internal models are continuously updated, then vocalization onset should progressively become lower and higher when participants receive FAF that is either increasing or decreasing, respectively. Thus, when participants receive unaltered auditory feedback at the end of the experiment, adaptation will be demonstrated if their F0 values begin near their F0 values from the final FAF trial. As a result, if singers rely more on an internal representation to control voice F0, then aftereffects should be more pronounced in singers than in nonsingers. Moreover, adaptation effects should generalize to a greater degree in singers than in nonsingers when asked to produce a different note (A4 and F4) with unaltered feedback following FAF trials.

II. METHODS

A. Participants

Twenty Wilfrid Laurier University students (all women) whose native tongue was North American English participated. Although there is no evidence to suggest that a gender difference exists in response to FAF, men were excluded so that all participants could adequately sing the same target notes. Of the 20 participants, 10 were trained singers (mean musical training was approximately 12 years) recruited from the faculty of music (vocal majors) at Wilfrid Laurier University. None of the trained singers reported having "perfect" pitch. The remaining ten participants were considered to be nonsingers, as none possessed any form of vocal training or ongoing participation in formal singing. All participants passed a bilateral hearing test at 20 dB for frequencies at 250, 500, 1000, 2000, 4000, and 8000 Hz and received financial compensation for their time and gave informed consent. The Wilfrid Laurier University Research Ethics Committee approved the procedures.

B. Apparatus

1. Participant recording sessions

Participants were situated in a double-walled sound attenuated booth (Industrial Acoustic Co., model 1601-01) and were fitted with headphones (Sennheiser HD 280 Pro) and a condenser microphone (Countryman Isomax E6 Omnidirectional Microphone), which was approximately 3 cm from their mouth. Multitalker babble noise (20 speakers simultaneously reading different passages; Auditec, St. Louis, MO) was presented at 80 dB SPL (sound pressure level) to limit natural acoustic feedback. Multitalker babble is the sound of 20 young adults simultaneously reading different passages, which is unintelligible to the listener. The target notes consisted of a female voice singing the consonant-vowel /ta/ that was presented at either 349, 392, or 440 Hz, respectively. Microphone signals were sent to a signal processor (VoiceOne 2.0, TC Helicon) that manipulated auditory feedback. The altered feedback was then mixed (Mackie ONYX 1640) with the multitalker babble and subsequently sent to the participant. Vocal productions were digitized at 44.1 kHz for future analysis.

2. Target stimuli recording

The stimuli used were the same as in Jones and Keough (2008). A trained singer produced the respective targets, F4, G4, and A4, which were processed using the speech modification algorithm STRAIGHT (speech transformation and representation using the adaptive interpolation of weighted spectrum; Kawahara *et al.*, 1999) to ensure that each target was exactly 349, 392 or 440 Hz.

C. Procedure

Participants matched a musical target over 210 trials, which were divided into three blocks of 70 trials. Each block consisted of 10 base line (only the last five trials were subjected to statistical analysis), 50 shift, and 10 test trials. During one block, participants reproduced the musical target G4 on all trials. Thus, participants received unaltered auditory feedback during the ten base line trials, followed by 50 FAF trials (2 cent increments to 100 cents, 1 semitone), and then ten more unaltered feedback trials. In two other blocks, participants reproduced either F4 or A4 for the ten base line and ten test trials while singing G4 during the FAF trials. Thus, participants only sang G4 during FAF trials. The other target notes, A4 and F4, were used to test whether participants exhibited sensorimotor adaptation when emulating notes that were never manipulated. Thus, the three block participants experienced were AGA, FGF, and GGG, where the first letter denotes the base line target (trials 1–10), the middle letter denotes the shifted target (trials 11–60), and the final letter denotes the test value (trials 61–70). The three blocks of trials were counterbalanced across participants.

On the first day of testing, participants produced the target on the three blocks (AGA, FGF, and GGG) of 70 trials and received auditory feedback that was shifted either upward or downward in 2 cent increments up to 100 or −100 cents, respectively, during the shift trials. Note that auditory feedback was shifted from the beginning of each utterance

until the end of their vocal productions. On a subsequent day, participants produced the remaining three blocks of 70 trials while receiving altered feedback that was shifted in the opposite direction. Thus, participants were required to produce the target for a total of six blocks of 70 trials that were counterbalanced over two days. Although aftereffects persisted for the duration of test trials in a previous study (Jones and Keough, 2008), there is no evidence in the literature to suggest that aftereffects continue to occur when tested on subsequent days. Individual trials commenced with the presentation of multitalker babble for 1000 ms followed by the target stimulus, which lasted 2000 ms in duration. Lastly, the multitalker babble was presented again for 4000 ms.

The babble was presented prior to the target to inform participants that the target note would be forthcoming. During the second presentation of the babble, participants were instructed to sing the target note as accurately as possible in pitch and duration (hold their pitch constant for approximately 2000 ms). Trials were initiated and controlled by a computer, and participants' vocal productions were recorded on a Macintosh G5 computer (TRACKTION v1, Woodinville, WA). F0 values for each vocal production were calculated, during offline analyses, using an autocorrelation algorithm included in the PRAAT program (Boersma, 2001). F0 values were normalized to each target note (F4, G4, or A4) by calculating the appropriate cent values using the following formula:

$$\text{Cents} = 100(12 \log_2 F/B)$$

where F is the F0 value in hertz and B is frequency of the target pitch participants were instructed to sing (349, 392, or 440 Hz).

The mean F0 values for singers and nonsingers during FAF trials (11–60) were calculated for each condition (AGA, GGG, and FGF) and are displayed in Figs. 1(a) and 1(b) (gray lines), respectively. Data for one nonsinger were removed from the statistical analyses as she exhibited poor F0 control during testing (mean F0 was 346 and 244 cents for the shifted up and down conditions, respectively). Only the first 1500 ms for each vocal production was analyzed because previous research has found that compensatory responses to FAF occur between 130 and 500 ms after perturbation onset (Burnett *et al.*, 1997; Burnett *et al.*, 1998; Jones and Munhall, 2002). The F0 values from the pitch shift up and down AGA, FGF, and GGG blocks were analyzed together and were broken into five different blocks of trials within each condition: shift trials (11–20, 21–30, 31–40, 41–50, and 51–60). Furthermore, the first five trials and the last five trials of each block were averaged and divided into early and late phases, respectively. Thus, a multivariate analysis of variance (MANOVA) was carried out on the mean F0 values with 2(experience:singer and nonsinger)×2(pitch shift:up and down)×5(block)×2(phase: early and late) as factors. Newman–Keuls' test was used for *post hoc* tests with an alpha level of 0.05 used for all statistical tests.

Also, the authors wanted to determine when singers' and nonsingers' mean F0 values during altered feedback trials were significantly different from base line F0 values. This test would indicate when compensation occurred in the
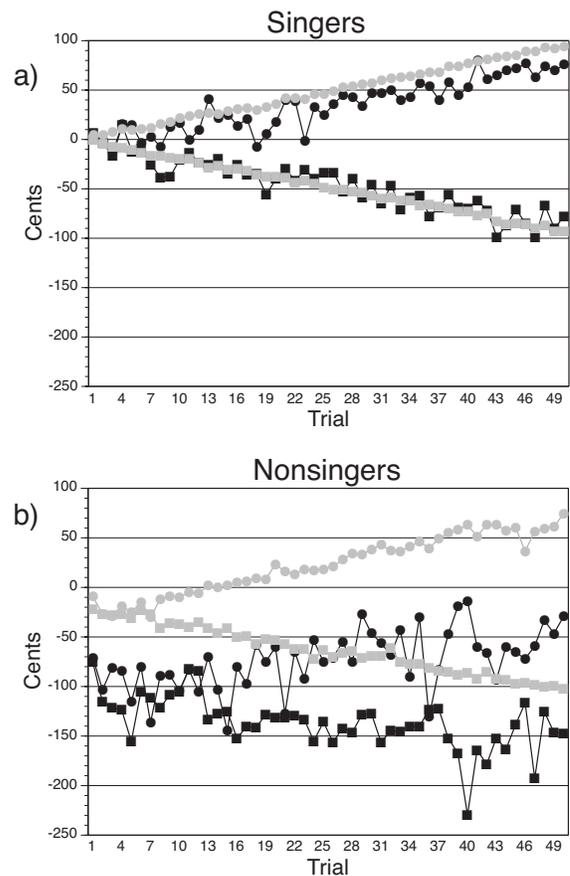

FIG. 1. (a) Singers' and (b) nonsingers' mean F0 (gray symbols) over a whole utterance and median F0 (black symbols) for the first 50 ms after vocal onset during FAF trials (11–60) for each condition (AGA, GGG, and FGF). Each data point represents the mean compensatory response to each incremental manipulation (three trials per shift value) averaged in blocks of 10 cent shifts during both upward and downward manipulations. Participants were required to match the musical target note G4 (392 Hz). The circles depict utterances that participants heard shifted downward; the squares depict utterances that participants heard shifted upward.

subtle manipulations in feedback. Furthermore, the authors examined the first test trial immediately after participants received altered feedback to determine if sensorimotor adaptation, in the form of aftereffects, occurred. In order to assess compensatory responses and aftereffects, multiple *t*-tests were performed on the mean F0 value for the last five base line trials compared with the mean F0 values on altered feedback trials and the first test trial during shifted up and down conditions.

Additionally, the median F0 value for the first 50 ms [see Figs. 1(a) and 1(b), black lines] of each utterance during the AGA, FGF, and GGG conditions was calculated for both singers and nonsingers across shifted trials. Determining whether differences occur between the initial 50 ms of initial shift trials and the later shifted trials indicated whether sensorimotor adaptation occurred during successive FAF trials. Furthermore, these median values were subtracted from the mean F0 over the entire utterance to determine the difference between where participants initiated vocal pitch and where they maintained their F0 values while singing. This test identified whether singers' and nonsingers' F0 productions were initiated at the desired target frequency or whether they per-

formed a searching strategy, increasing or decreasing their F0 to match the target. The median values for the first 50 ms and the median less the mean values were then categorized into five blocks of trials and further divided into two phases, in the same fashion as the mean F0 values. Thus, two separate MANOVAs were performed on each of the aforementioned values (median and median minus mean) during the shifted up and down AGA, FGF, and GGG conditions with 2(experience:singer and nonsinger)×2(pitch shift:up and down)×5(block)×2(phase:early and late) as factors.

Finally, in order to elucidate when or if sensorimotor adaptation commences during FAF trials and whether aftereffects persist during test trials, multiple t-tests (Bonferroni corrected) were conducted on the median F0 values for the first 50 ms during both the shifted up and down conditions. The average of the median F0 values for the last five base line trials was compared with the median F0 values of all shift trials and the first test trial following exposure to FAF.

III. RESULTS

The MANOVA carried out on the mean F0 values with 2(experience: singer and nonsinger)×2(pitch shift:up and down)×5(block)×2(phase:early and late) as factors revealed a main effect of experience and pitch shift, $F(1,55) = 12.59, p < 0.05$ and $F(1,55) = 634.73, p < 0.05$, respectively. Nonsingers' mean F0 values during FAF trials were found to be significantly lower than the pitch target (392 Hz, G4) compared to the mean F0 values of singers [see Figs. 1(a) and 1(b)]. Also, the mean F0 values during the pitch shift up condition were found to be lower than the F0 values during the pitch shift down condition. A significant two-way interaction was found between pitch shift and block, $F(4,220) = 423.40, p < 0.05$. *Post hoc* analysis revealed that the first block (trials 11–20) of mean F0 values during the pitch shift up condition and the pitch shift down condition was significantly different (progressively lower and higher during pitch shift up and down conditions, respectively) from all other blocks of pitch shift F0 values obtained during FAF trials (21–60) ($p < 0.05$) [see Figs. 1(a) and 1(b)]. Note that each data point in Fig. 1 represents the mean compensatory response to each incremental manipulation (three trials per shift value; one from each of AGA, GGG, and FGF conditions) averaged in blocks of 10 cent shifts during both upward and downward manipulations. This pattern demonstrates that participants were compensating more during later trials to increasing and decreasing shifts in auditory feedback. Interestingly, this suggests that both singers and nonsingers initiate compensatory responses quite early to subtle changes in auditory feedback. Furthermore, the mean F0 values obtained for each pitch shift block during the shift up condition were significantly lower than the mean F0 values for each block in the shift down condition ($p < 0.05$).

A two-way interaction between pitch shift and phase was also found to be significant, $F(1,55) = 121.05, p < 0.05$. *Post hoc* analysis indicated that the mean F0 values during the first phase of the shift up condition and the shift down condition were significantly different from the mean F0 values during the second phase of the shift up and down condi-

tions, respectively ($p < 0.05$). This implies that as pitch shifts progressively increased or decreased on FAF trials participants correspondingly adjusted their F0 to continue to produce the target notes accurately, and given that the late phase shift trials were larger than early phase shift trials it is not surprising that the mean F0 values were found to differ, on average, between the two phases. Also, the F0 values during the first and second phases of the shift up condition differed significantly from the F0 values of the first and second phases of the shift down condition ($p < 0.05$). No other significant main effects or interactions were observed.

Multiple t-tests (uncorrected) were carried out on the mean F0 during the shifted up and down conditions for singers and nonsingers. When three consecutive significant differences were found during the shifted trials, the first significant response from base line was used to indicate the initial compensatory response. During the shift up trials, when the mean of the last five base line trials was compared with the mean of the shift trials it was found that singers initiated compensatory responses on the third (6 cents) shift trial, $t(60) = -2.89, p < 0.05$. The same initial response was also observed during the shift down condition, $t(60) = 2.67, p < 0.05$. Moreover, singers' F0 values on shift up and down trials remained significantly different from their base line F0 values. Thus, singers' F0 values on shifted trials were significantly different from the average of the last five base line trials on 48 of 50 trials during both shift up and down conditions.

On the other hand, during the shift up condition, nonsingers were found to initiate compensatory responses on shift trial 13 (26 cents), $t(60) = -3.68, p < 0.05$, with the remaining F0 values remaining significantly different from the base line F0 values. The only difference between the shift up and down conditions for nonsingers was that they initiated compensatory responses two trials earlier during the down condition, at shift trial 11 (22 cents), $t(60) = 3.12, p < 0.05$. Thus, nonsingers' F0 values were different from base line on 38/50 and 40/50 during the shift up and down conditions, respectively.

The authors were also interested in determining whether compensating for FAF would result in aftereffects (for the GGG condition) when auditory feedback was returned to normal and whether these aftereffects would generalize (for the AGA and FGF conditions) to a note other than the one participants received during testing. Thus, t-tests were carried out on the average of the mean F0 values of the last five base line trials for the AGA, FGF, and GGG conditions and the first test trial (when auditory feedback was returned to normal following FAF trials) of each respective condition [see Figs. 2(a) and 2(b)]. Nonsingers' average base line F0 values for the GGG trials were significantly different from the F0 values for the initial test trial for both shift down and up conditions, $t(8) = -2.90, p < 0.05$ and $t(8) = 2.30, p < 0.05$, respectively. Moreover, singers' base line mean F0 values for both shift down and up conditions were also significantly different from their initial mean F0 values during testing, $t(9) = -5.39, p < 0.05$ and $t(9) = 4.49, p < 0.05$, respectively. Thus, both singers and nonsingers exhibited aftereffects following FAF trials. That is, their mean F0 values were above
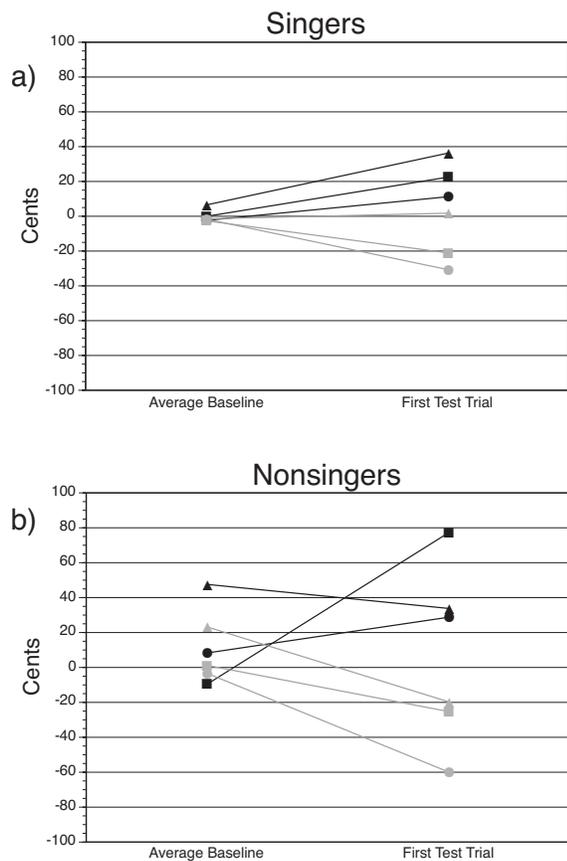
FIG. 2. The mean F0 of utterances produced by (a) singers and (b) nonsingers, averaged over the last five base line trials (prior to receiving FAF), and the mean F0 for the first test trial utterance following FAF. The triangles represent F0 values for the target note F4 (349 Hz), the squares represent F0 values for the target note G4 (392 Hz), and the circles represent F0 values for the target note A4 (440 Hz). Gray symbols represent F0 values obtained during the upward shift condition, and the black symbols represent F0 values obtained during the downward shift condition.

and below the target F0 when auditory feedback was returned to normal following exposure to upward and downward FAFs.

When participants sang a different note following FAF trials (either F4 or A4), nonsingers' mean F0 values were significantly different from their average base line F0 values for the shifted up AGA condition, $t(8)=2.58, p<0.05$. On the other hand, singers' mean base line F0 values were significantly different from their mean test F0 values for both the AGA and FGF shifted down conditions, $t(9)=-2.35, p<0.05$ and $-3.72, p<0.05$, respectively, and the AGA shifted up condition, $t(9)=5.51, p<0.05$. Overall, both groups exhibited aftereffects that generalized to at least one note other than the one produced during FAF trials. However, when the alpha level was corrected (Bonferroni) for multiple t-tests, nonsingers' mean F0 base line values failed to remain statistically significant when compared to the initial test trial for all conditions. Singers' aftereffects during GGG trials remained significantly different during both shifted down and up conditions, but the only generalization effect to remain significant were the differences in F0 values obtained during the shifted up AGA base line and test conditions.

A MANOVA was performed on the median 50 ms F0 values (the median F0 value within the first 50 ms after vocal onset) during the shifted up and down AGA, FGF, and GGG conditions with 2(experience: singer and nonsinger) $\times$ 2(pitch shift: up and down) $\times$ 5(block) $\times$ 2(phase: early and late) as factors. The data obtained during testing revealed a significant main effect of experience and pitch shift condition, $F(1,55)=15.11, p<0.05$ and $F(1,55)=57.72, p<0.05$, respectively. Nonsingers' median F0 values were determined to be significantly lower than singers' median F0 values. The median F0 values during the shift down condition were found to be significantly higher than the median F0 values obtained during the shift up condition. A two-way interaction between pitch shift and block was significant, $F(4,220)=45.58, p<0.05$. Post hoc testing indicated that the median F0 values during the first block of shift trials on the shift up condition were significantly higher than all remaining shift blocks ($p<0.05$). Moreover, the median F0 values during the first block of shift trials during the shift down condition were significantly lower than shift blocks 3, 4, and 5 ($p<0.05$). Also, the median F0 values for all shift up blocks were significantly lower than the median F0 values for all shift down blocks ($p<0.05$).

Additionally, there was a significant interaction between pitch shift and phase, $F(1,55)=7.39, p<0.05$. Post hoc testing revealed that there were no differences between the early and late phases of the shift up and down conditions, respectively ($p>0.05$). However, the median F0 values of the early and late phases of the shift down condition were significantly higher than the median F0 values of both the early and late phases of the shift up condition ($p<0.05$). Finally, there was a significant three-way interaction between experience, pitch shift, and block, $F(4,220)=3.01, p<0.05$. Post hoc analysis revealed that nonsingers' first block of median F0 values during the initial block of the shift down condition was significantly lower than the median F0 values during blocks 3, 4, and 5 ($p<0.05$). Also, during the shift up condition, nonsingers' median F0 values during the initial shift block were significantly higher than the median F0 values on blocks 3, 4, and 5 ($p<0.05$). Singers' initial shift block median F0 values for both the shift down and up conditions were significantly higher and lower, respectively, than the median F0 values on shift blocks 3, 4, and 5 ($p<0.05$). Moreover, singers' median F0 values on shift down trials were significantly higher than all nonsingers' median F0 values on shift down trials ($p<0.05$). Singers' median F0 values on shift up trials were also found to be significantly higher than all nonsingers' median F0 values on shift up trials ($p<0.05$). No other significant main effects or interactions were observed.

Multiple t-tests (uncorrected) were also performed on the median values of the first 50 ms of each utterance. Similar to the multiple t-tests reported previously, three consecutive significant differences were required prior to establishing when adaptation occurred. The last five base line median F0 values were averaged to establish a base line for comparison. This value was then compared to all shifted and test trials. During the shifted up condition, singers' median F0 values were found to be different from base line at shift trial 17 (34 cents), $t(60)=-2.59, p<0.05$. In total, singers' me-

dian F0 values were different from base line on 33/50 shifted up trials. During the shifted down condition, singers exhibited adaptation responses at shift trial 13 (26 cents), $t(60) = 3.94, p < 0.05$. During the shifted down condition, singers' median F0 values were different than base line on 35/50 trials. Interestingly, nonsingers' median F0 values failed to significantly differ on three consecutive occasions during the shifted up or down conditions. However, nonsingers' median F0 values did differ on two consecutive trials during both the shifted up and down conditions (but only once during each). Nonsingers' median F0 values differed at shift trial 38 (76 cents) and 39 (78 cents), $t(60) = -2.73$ and 2.65, $p < 0.05$, during shifted up and down conditions, respectively. Furthermore, nonsingers' median F0 values only varied from base line on 12/50 and 4/50 during the shifted up and down conditions, respectively.

Lastly, a MANOVA was performed on the median (50 ms F0) minus the mean (1500 ms) F0 values obtained during the shifted up and down AGA, FGF, and GGG conditions with 2(experience:singer and nonsinger)×2(pitchshift:up and down)×5(block)×2(phase:early and late) as factors. When the median F0 values for the first 50 ms of the shifted down and up conditions were subtracted from the mean F0 values for each respective condition, results indicated that there was a significant main effect of experience, $F(1, 55) = 12.52, p < 0.05$. Nonsingers' F0 values were found to be significantly lower than the F0 values for singers. A main effect of pitch shift was also found to be significant, $F(1, 55) = 6.90, p < 0.05$. F0 values during the pitch shift down condition were found to be significantly higher than the F0 values during the pitch shift up condition. Results identified a significant two-way interaction between pitch shift and block, $F(4, 220) = 6.86, p < 0.05$. *Post hoc* testing indicated that the only difference during the pitch shift down condition was that the initial block of F0 values was significantly lower than the last block of shift trials ($p < 0.05$). Moreover, the initial shift block F0 values during the shift down condition were significantly different from the F0 values of shift blocks 4 and 5 of the shift up condition ($p < 0.05$). The initial block of shift trials during the shift up condition was determined to be significantly higher than shift blocks 4 and 5 ($p < 0.05$), as well as the last block of trials during the shift down condition ($p < 0.05$).

Finally, there was a significant three-way interaction between experience, pitch shift, and block, $F(4, 220) = 2.72, p < 0.05$. *Post hoc* testing revealed that nonsingers' first block of F0 values during the shift down condition was significantly lower than the F0 values during the last block of shift trials ($p < 0.05$). Nonsingers' F0 values during the initial block of shift up trials were significantly higher than the F0 values of blocks 4 and 5 ($p < 0.05$). Singers' F0 values during the initial block of shift down trials were not significantly different from any other block of shift down trials ($p > 0.05$). The F0 values of the initial block of shift up trials were also not different from any other block of shift up trials ($p > 0.05$). Furthermore, nonsingers' F0 values during the shift up and down conditions were all significantly lower than singers' F0 values ($p < 0.05$). No other significant main effects or interactions were observed.

IV. DISCUSSION

The purpose of this study was to investigate the sensitivity of the mechanisms that modify internal representations in singers and nonsingers when presented with subtle changes in auditory feedback while singing. In accord with our predictions, singers and nonsingers use auditory feedback to compensate for subtle manipulations in auditory feedback while singing. That is, both singers and nonsingers compensated for the altered feedback by increasing or decreasing their F0 to downward and upward shifts in feedback. The main difference between the groups was that singers are immediately and consistently more accurate when they match the target notes. Nonsingers' F0 values were consistently lower than the pitch target, regardless of the direction of the manipulation. An important observation was that singers initiated compensatory responses to altered feedback on the third shift (6 cents) trial during both upward and downward altered conditions. On the other hand, nonsingers initiated compensatory responding at shift trials 13 (26 cents) and 11 (22 cents) during upward and downward manipulations, respectively.

Furthermore, the authors examined the first 50 ms of each vocal production in order to determine how participants initiate F0 responses while singing. If participants adapt to the altered feedback by altering an internal representation that regulates their initial F0 production, their initial F0 production should be close to the F0 values produced while compensating during the previous trial. Results showed that singers, and to a lesser degree nonsingers, compensated for gradual FAF manipulations by starting subsequent utterances at similar F0 values obtained on the previous FAF trial. Thus, singers were continually updating their internal model to account for the consistently increasing or decreasing changes in F0 so they could initiate voice F0 at the desired target frequency. On the other hand, nonsingers appeared to search for the target note by starting below the auditory target of 0 cents and by increasing their F0 until they matched the note. Moreover, multiple *t*-tests failed to provide evidence to suggest that sensorimotor adaptation occurred in nonsingers during the altered feedback conditions. In the case of singers, multiple *t*-tests on the median 50 ms F0 data revealed that sensorimotor adaptation occurred at 34 and 26 cents during shifted up and down conditions, respectively.

Previous research using the FAF paradigm has typically examined vowel phonation (Burnett *et al.*, 1997; Burnett *et al.*, 1998; Burnett and Larson, 2002; Elman, 1981; Jones and Munhall, 2000, 2002, 2005; Natke *et al.*, 2003; Toyomura *et al.*, 2007), with an emphasis on compensatory responding. Not surprisingly, the results are consistent with the data collected on singing (Burnett *et al.*, 1997; Jones and Keough, 2008; Natke *et al.*, 2003, Zarate and Zatorre, 2005, 2008) in that participants compensate for manipulations in auditory feedback by increasing or decreasing their F0 in the opposite direction of the perturbation. The data obtained from the current experiment are in accord with this observation, as both singers and nonsingers altered their F0 to similar degrees while receiving altered feedback.

It has been argued that the aftereffects observed after exposure to FAF are a result of modifying an acoustic-motor representation (Jones and Munhall, 2000, 2002; Jones and Keough, 2008; Natke *et al.*, 2003). However, sensorimotor adaptation has not been extensively examined in singers. When Jones and Keough (2008) returned trained singers' auditory feedback to normal, following FAF trials, their F0 values never returned to base line levels. Thus, it appears that brief exposure to altered feedback resulted in a remapping of an internal model (Jones and Keough, 2008). In the current study, the authors also found a global remapping to subtle changes in auditory feedback. Uncorrected multiple *t*-tests revealed that aftereffects existed for both singers and nonsingers. That is, when auditory feedback was suddenly returned to normal, following exposure to FAF, singers' and nonsingers' mean test trial F0 values were significantly higher and lower (following shifted down and up feedback, respectively) than base line.

These aftereffects also generalized to a note other than the one they produced during FAF trials. For instance, singers' mean F0 values for the target note A4 during shift up and down conditions and F4 following shift down trials were significantly different from the average of the last five base line trials. Nonsingers' test F0 values were only different from base line for the target note A4 during the shift up feedback condition. Thus, participants modified an internal representation that regulates F0 control. As a consequence, the pattern of responding observed in the current study is similar to the pattern of responding that has been found in previous FAF studies (Jones and Munhall, 2000, 2002; Jones and Keough, 2008). However, when the alpha level was corrected for the multiple *t*-tests, only the aftereffects observed for singers and the generalization for the target note A4 during the shifted up condition remained significant.

Previous studies have identified aftereffects following FAF but have only examined adaptation after a period of training at the end of an experiment (Jones and Munhall, 2000, 2002; Jones and Keough, 2008). For this study, the authors tracked F0 values early in each utterance during the training period to determine whether singers and nonsingers continuously adjust their internal motor representation for F0 control in response to changes in feedback. Singers were found to match the target, on average, more accurately than nonsingers. That is, they initiated their vocal productions near the F0 frequencies that were required to compensate for the FAF experienced in previous trials. This sensorimotor adaptation in singers occurred when the discrepancy between expected F0 output and auditory feedback was approximately 30 cents. Thus, subtle discrepancies in feedback can be accounted for by the mechanisms that support singers' internal model for vocal control. On the other hand, nonsingers consistently started their productions below the target. Their F0 values began consistently below the pitch target and were adjusted (upward) until they reached the note, as best as they could. This difference between singers and nonsingers is consistent with the data obtained in Jones and Keough (2008).

One difference found in previous work investigating sustained vowel phonation and singing using FAF was the level of compensation observed. Typically, it has been reported that compensatory responses do not exceed 65 cents for shifts up to 600 cents. However, the bulk of previous research has focused on sustained vowel production while receiving random pitch perturbations (e.g., Burnett *et al.*, 1997; Donath *et al.*, 2002; Larson, 1998; Larson *et al.*, 2000; Liu and Larson, 2007). Jones and Keough (2008) reported that nonsingers fully compensated (100 cents) for highly predictable shifts in FAF almost immediately, and although singers initially showed partial compensation (approximately 65 cents) they eventually altered their F0 values to accommodate for the altered feedback. Data from the current study revealed that singers and nonsingers exhibited near perfect levels of compensation during predictable shifted up and down conditions. Thus, it appears that auditory feedback may be used in a task-dependent manner such that when achieving a particular pitch target is important, as in singing, auditory feedback guides production.

Although singers and nonsingers eventually compensated for the FAF to the same degree, the point at which each group altered their productions based on the auditory feedback differed. Singers began compensating when they detect feedback errors as large as 6 cents (1.36 Hz) during both the shift up and down conditions. This value is consistent with Sundberg's (1987) finding that trained singers can correct for production errors with an accuracy of less than 1 Hz from an intended pitch target (A4, 440 Hz). On the other hand, nonsingers initiated compensatory responses at approximately 24 cents (5.47 Hz). Data from both groups fall within the just-noticeable difference range found by Pape and Mooshammer (2006) of F0 contours for natural stimuli. Moreover, nonsingers' values are also very similar to the average threshold reported by Hafke (2008), who found that pitch shift changes were not reliably identified when they were below 26 cents. However, Loui *et al.* (2008) found that control participants' psychophysical thresholds of perception and production were around 2.0 and 2.5 Hz, respectively. Although the authors never examined the perceptual aspect directly, the data they obtained for singers and nonsingers are relatively consistent with the production threshold findings of Loui *et al.* (2008).

Based on the current findings, the authors believe that singers, due to their extensive training and experience, are more capable of compensating for subtle manipulations of auditory feedback earlier than nonsingers. Moreover, singers more readily alter their internal representations to prevent the occurrence of these errors in subsequent utterances. On the other hand, when singers detect large incongruities between perception and production they rely more on their internal model to produce the target (Jones and Keough, 2008). Singers may have developed different internal models during vocal training that provides them with a greater capacity and flexibility to control voice F0 during different task demands. The F0 control system may deem certain feedback discrepancies to be too large to be internally generated. Zarate and Zatorre (2005, 2008) found similar results when singers and nonsingers were exposed to FAF. Zarate and Zatorre (2005, 2008) asked participants to ignore the feedback and continually reproduce the target as accurately as possible. The pat-

tern of behavioral results suggests that singers could successfully ignore their altered (200 cents) feedback and continue to produce the target notes accurately. Interestingly, nonsingers were unable to ignore the feedback as their F0 responses indicated partial compensation. These findings suggest that singers' internal models are flexible in nature; the models can be adjusted to compensate for perturbations as small as 6 cents (1.36 Hz) when instructed to match the target note, or the models can be adjusted to ignore pitch errors (e.g., 200 cents) when instructed to do so by relying on the existing motor representation (Zarate and Zatorre, 2008).

Given that behavioral differences exist between singers and nonsingers during FAF, differences in neural activity between the two groups may be expected. Zarate and Zatorre (2005, 2008) found that despite differences in vocal production accuracy during normal feedback conditions, both singers and nonsingers exhibited similar functional networks for singing. These areas included the bilateral auditory cortices, bilateral primary motor cortices, the supplementary motor area (SMA), the anterior cingulate cortex (ACC), thalamus, insula, and cerebellum. These results are consistent with a previous study carried out in the Zatorre laboratory (Perry et al., 1999) and by Toyomura et al. (2007). However, when exposed to FAF and asked to ignore the FAF, singers showed enhanced activation in the inferior parietal lobule (IPL), superior temporal gyrus (STG), superior temporal sulcus (STS), and right insula (Zarate and Zatorre, 2005). On the other hand, enhanced activation in the ACC, STS, insula, putamen, pre-SMA, and IPL was observed in singers when they were directly asked to compensate for the FAF (Zarate and Zatorre, 2005). The authors conclude that the additional activation of the STG and the STS in singers during FAF conditions is suggestive of an increased perceptual analysis of the incoming signal (Zarate and Zatorre, 2005). In addition, the authors point out that an increased activation in the putamen of singers during both ignore and compensate conditions suggests that singers are relying on well-defined internal representations to sing the targets while receiving FAF (Zarate and Zatorre, 2008). Indeed, it may be the case that singers rely on more than one internal model to regulate voice F0.

V. Conclusion

Overall, results showed that singers and nonsingers compensated for FAF to a similar degree. On the other hand, singers were more accurate in their pitch productions, as their F0 values were consistently closer to the intended pitch target. Singers compensated for FAF when auditory feedback was manipulated up or down by 6 cents, whereas nonsingers compensated when feedback was shifted upward by 26 cents and downward by 22 cents. Additionally, examining the first 50 ms of vocal productions during FAF trials showed that singers continued to initiate F0 values near the intended target frequency. The authors argue that this resulted from the online recalibration of an internal model regulating voice F0. Nonsingers were found to consistently initiate F0 productions below the pitch target and to increase their F0 until they matched the note. Thus, singers and nonsingers rely on an internal model to regulate voice F0, but singers' models appear to be more sensitive in response to subtle discrepancies in auditory feedback.

It appears that using absolute targets in a predictable FAF paradigm permitted a unique examination of F0 control. For instance, examining sensorimotor adaptation during singing has proven to be informative in testing the acoustic-motor mapping of F0 control in musically untrained and trained individuals. Using a specialized subgroup of the population may allow us to better understand how sensitive internal models are to FAF and how these motor commands are represented neurologically. Also, it is arguable that voice F0, while singing, is represented by multiple internal models that correspond to specific musical notes. Indeed, current research in our laboratory examines how sensory feedback can be used to continuously recalibrate multiple internal models regulating F0 while singing.

Bauer, J. J., Mittal, J., Larson, C. R., and Hain, T. C. (**2006**). "Vocal responses to unanticipated perturbations in voice loudness feedback: An automatic mechanism for stabilizing voice amplitude," J. Acoust. Soc. Am. **119**, 2363–2371.

Boersma, P. (**2001**). "Praat, a system for doing phonetics by computer," Glot International **5**, 341–345.

Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (**1998**). "Voice F0 responses to manipulations in pitch feedback," J. Acoust. Soc. Am. **103**, 3153–3156.

Burnett, T. A., and Larson, C. R. (**2002**). "Early pitch-shift response is active in both steady and dynamic voice pitch control," J. Acoust. Soc. Am. **112**, 1058–1063.

Burnett, T. A., Senner, J. E., and Larson, C. R. (**1997**). "Voice F0 responses to pitch-shifted auditory feedback: A preliminary study," J. Voice **11**, 202–211.

Donath, T. M., Natke, U., and Kalveram, K. T. (**2002**). "Effects of frequency-shifted auditory feedback on voice F0 contours in syllables," J. Acoust. Soc. Am. **111**, 357–366.

Elman, J. L. (**1981**). "Effects of frequency-shifted feedback on the pitch of vocal productions," J. Acoust. Soc. Am. **70**, 45–50.

Fairbanks, G. (**1954**). "Systematic research in experimental phonetics. I. A theory of the speech mechanism as a servosystem," J. Speech Hear. Disord. **19**, 133–139.

Finney, S., and Palmer, C. (**2003**). "Auditory feedback and memory for music performance: Sound evidence for an encoding effect," Mem. Cognit. **31**, 51–64.

Hafke, H. Z. (**2008**). "Nonconscious control of fundamental voice frequency," J. Acoust. Soc. Am. **123**, 273–278.

Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., and Kenney, M. K. (**2000**). "Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex," Exp. Brain Res. **130**, 133–141.

Houde, J. F., and Jordan, M. I. (**1998**). "Sensorimotor adaptation in speech production," Science **279**, 1213–1216.

Jones, J. A., and Keough, D. (**2008**). "Auditory-motor mapping for pitch control in singers and nonsingers," Exp. Brain Res. **190**, 279–287.

Jones, J. A., and Munhall, K. G. (**2000**). "Perceptual calibration of F0 production: Evidence from feedback perturbation," J. Acoust. Soc. Am. **108**, 1246–1251.

Jones, J. A., and Munhall, K. G. (**2002**). "The role of auditory feedback during phonation: Studies of Mandarin tone production," J. Phonetics **30**, 303–320.

Jones, J. A., and Munhall, K. G. (**2005**). "Remapping auditory-motor representations in voice production," Curr. Biol. **15**, 1768–1772.

Kawahara, H. (**1998**). In *Computational Auditory Scene Analysis*, edited by D. F. Rosenthal and H. G. Okuno (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 335–349.

Kawahara, H., Masuda-Katsuse, I., and de Cheveigne, A. (**1999**). "Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds," Speech Commun. **27**, 187–207.

Kirchner, J. A., and Wyke, B. D. (**1965**). "Articular reflex mechanisms in the larynx," Ann. Otol. Rhinol. Laryngol. **74**, 749–768.

Lane, H., and Tranel, B. (**1971**). "The Lombard sign and the role of hearing in speech," J. Speech Hear. Res. **14**, 677–709.

Larson, C. R. (**1998**). "Cross-modality influences in speech motor control: The use of pitch shifting for the study of f0 control," J. Commun. Disord. **31**, 489–503.

Larson, C. R., Altman, K. W., Liu, H., and Hain, T. C. (**2008**). "Interactions between auditory and somatosensory feedback for voice F0 control," Exp. Brain Res. **187**, 613–621.

Larson, C. R., Burnett, T. A., Kiran, S., and Hain, T. C. (**2000**). "Effects of pitch-shift velocity on voice F0 responses," J. Acoust. Soc. Am. **107**, 559–564.

Lee, B. S. (**1950**). "Effects of delayed speech feedback," J. Acoust. Soc. Am. **22**, 824–826.

Liu, H., and Larson, C. R. (**2007**). "Effects of perturbation magnitude and voice F level on the pitch-shift reflex," J. Acoust. Soc. Am. **122**, 3671–3677.

Loui, P., Guenther, F. H., Mathys, C., and Schlaug, G. (**2008**). "Action-perception mismatch in tone-deafness," Curr. Biol. **18**, R331–R332.

Natke, U., Donath, T. M., and Kalveram, K. T. (**2003**). "Control of voice fundamental frequency in speaking versus singing," J. Acoust. Soc. Am. **113**, 1587–1593.

Oller, D. K., and Eilers, R. E. (**1988**). "The role of audition in infant babbling," Child Dev. **59**, 441–449.

Pape, D., and Mooshammer, C. (**2006**). "Is intrinsic pitch language dependent? Evidence from a cross-linguistic vowel pitch experiment (with additional screening of the listeners' DL for music and speech)," in MULTILING-2006, Paper No. 018.

Perry, D. W., Zatorre, R. J., Petrides, M., Alivisatos, B., Meyer, E., and Evans, A. (**1999**). "Localization of cerebral activity during simple singing," NeuroReport **10**, 3979–3984.

Proteau, L., Marteniuk, R. G., Girouard, Y., and Dugas, C. (**1987**). "On the type of information used to control and learn an aiming movement after moderate and extensive training," Hum. Mov. Sci. **6**, 181–199.

Proteau, L., Marteniuk, R. G., and Levesque, L. (**1992**). "A sensorimotor basis for motor learning: Evidence indicating specificity of practice," Q. J. Exp. Psychol. **44A**, 557–575.

Purcell, D. W., and Munhall, K. G. (**2006**). "Compensation following real-time manipulation of formants in isolated vowels," J. Acoust. Soc. Am. **119**, 2288–2297.

Sapir, S., McClean, M. D., and Larson, C. R. (**1983**). "Human laryngeal responses to auditory stimulation," J. Acoust. Soc. Am. **73**, 315–321.

Sundberg, J. (**1987**). *The Science of the Singing Voice* (Northern Illinois University Press, Dekalb, IL).

Toyomura, A., Koyama, S., Miyamoto, T., Terao, A., Omori, T., Murohashi, H., and Kuriki, S. (**2007**). "Neural correlates of auditory feedback control in human," Neuroscience **146**, 499–503.

Watts, C., Murphy, J., and Barnes-Burroughs, K. (**2003**). "Pitch-matching accuracy of trained singers, untrained subjects with talented singing voices, and untrained subjects with non-talented singing voices in conditions of varying feedback," J. Voice **17**, 187–196.

Wyke, B. D. (**1974**). "Laryngeal neuromuscular control systems in singing: A review of current concepts," Folia Phoniatr (Basel) **26**, 295–306.

Yates, A. J. (**1963**). "Delayed auditory feedback," Psychol. Bull. **60**, 213–251.

Yoshida, Y., Saito, T., Hirano, M., and Kanaseki, T. (**1989**). "Laryngeal sensory innervation: Origins of sensory nerve fibers in the nodose ganglion of the cat," J. Voice **3**, 314–320.

Zarate, J. M., and Zatorre, R. J. (**2005**). "Neural substrates governing audio-vocal integration for vocal pitch regulation in singing," Ann. N.Y. Acad. Sci. **1060**, 404–408.

Zarate, J. M., and Zatorre, R. J. (**2008**). "Experience-dependent neural substrates involved in vocal pitch regulation during singing," Neuroimage **40**, 1871–1887.