Laryngeal electromyographic responses to perturbations in voice pitch auditory feedback

Hanjun Liu
Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-Sen University, Guangzhou, 510080, People's Republic of China

Roozbeh Behroozmand
Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois 60208

Michel Bove
Department of Otolaryngology, Head and Neck Surgery, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611

Charles R Larson[a]
Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois 60208

The present study was conducted to test the hypothesis that intrinsic laryngeal muscles are involved in producing voice fundamental frequency (F_0) responses to perturbations in voice pitch auditory feedback. Electromyography (EMG) recordings of the cricothyroid and thyroarytenoid muscles were made with hooked-wire electrodes, while subjects sustained vowel phonations at three different voice F_0 levels (conversational, high pitch in head register, and falsetto register) and received randomized pitch shifts (± 100 or ± 300 cents) in their voice auditory feedback. The median latencies from stimulus onset to the peak in the EMG and voice F_0 responses were 167 and 224 ms, respectively. Among the three different F_0 levels, the falsetto register produced compensatory EMG responses that occurred prior to vocal responses and increased along with rising voice F_0 responses and decreased for falling F_0 responses. For the conversational and high voice levels, the EMG response timing was more variable than in the falsetto voice, and changes in EMG activity with relevance to the vocal responses did not follow the consistent trend observed in the falsetto condition. The data from the falsetto condition suggest that both the cricothyroid and thyroarytenoid muscles are involved in generating the compensatory vocal responses to pitch-shifted voice feedback.
© 2011 Acoustical Society of America. [DOI: 10.1121/1.3575593]

I. INTRODUCTION

The control of voice fundamental frequency (F_0) involves several central neural mechanisms including operations such as pitch memory recall and comparison with auditory feedback for the planning and execution of vocal tasks. At a more peripheral level, groups of muscles determine voice F_0 by controlling the length, tension, and three-dimensional geometry of the vocal folds. Muscles of the thorax and abdomen are also involved in the regulation of air pressure and flow. The comparative ease of the study of laryngeal muscles involved in voice control has fostered many studies and has led to a fairly good understanding of their functions (Hirano and Ohala, 1969; Hirose and Gay, 1972; Lindestad *et al.*, 1991; Chanaud and Ludlow, 1992; Ludlow and Lou, 1996; Barkmeier *et al.*, 2000; Loucks *et al.*, 2005; Luschei *et al.*, 2006a). By contrast, studies of central neural mechanisms have been difficult, and our understanding of how the brain controls the voice is rather poor.

One technique for the study of central neural mechanisms of voice F_0 control that has been widely used is that of auditory perturbation, wherein voice pitch feedback is unexpectedly altered as subjects are vocalizing or speaking (Elman, 1981; Kawahara, 1995; Burnett *et al.*, 1998; Jones and Munhall, 2000; Bauer *et al.*, 2006). It has been demonstrated that voice F_0 responses to pitch-shifted feedback are mostly opposite in direction to the perturbation. That is, people lower their voice F_0 when their voice pitch feedback is shifted upward, and they raise their F_0 for downward stimuli. Such compensatory responses are thought to act as part of a negative feedback controller system that helps to stabilize F_0 (Hain *et al.*, 2000). It has been suggested that the responses are automatic or reflexive in nature because subjects seem to be unaware that they are producing rapid (about 100 ms latency) and direction-sensitive vocal responses to feedback perturbation (Hain *et al.*, 2000). Thus, the audio–vocal system relies on auditory feedback to produce changes in voice F_0 that automatically correct for errors in vocal production.

In more recent studies, it has also been found that the pitch-shift reflex (PSR) can be modulated by specific vocal tasks. For example, when pitch-shifted voice feedback is presented to English speakers during speech transitions,

[a]Author to whom correspondence should be addressed. Electronic mail: clarson@northwesern.edu

vocal responses can be larger and quicker than responses during a sustained vowel phonation (Chen *et al.*, 2007). Greater response magnitudes were also produced while singing compared to speaking (Natke *et al.*, 2003). These observations suggest that the neural mechanisms of the PSR can be modulated according to the specific demands of the vocal task.

With our growing understanding of the role of auditory feedback for the control of voice F_0, it is also important to understand how corrective vocal motor commands from the central nervous system are implemented at the peripheral level. Over the past 40–50 years, several studies have been done to define the role of individual muscles in the control of voice F_0 during voluntary vocalizations (Hirano and Ohala, 1969; Hirano *et al.*, 1970; Hirose and Gay, 1972; Larson *et al.*, 1987; Lindestad *et al.*, 1990; 1991; Luschei *et al.*, 1999; Luschei *et al.*, 2006b; Sataloff *et al.*, 2006). Of the intrinsic laryngeal muscles, the cricothyroid (CT) is thought to be one of the most important because it regulates the length of the vocal folds and thereby the stiffness of the vocal fold epithelium. Electromyography (EMG) recordings of the CT have shown that it increases and decreases its activity in a direct relation with corresponding changes in voice F_0 (Hirano *et al.*, 1970). The vocalis muscle affects the internal stiffness of the body of the vocal folds and shows changes in EMG levels with corresponding changes in F_0, but unlike the CT, it ceases its activity with a transition into the falsetto register (Hirano *et al.*, 1970). The thyromuscularis portion of the thyroarytenoid (TA) also seems to have a direct role in the control of voice F_0, and it is involved in many other functions such as swallowing and glottal adduction (Titze, 1994; Sataloff *et al.*, 2006). Other muscles, such as the lateral cricoarytenoid, interarytenoid and posterior cricoarytenoids appear to have a greater role in regulating the dimensions of the glottis, vocal fold opening and closure, and their role in the regulation of F_0 is less significant than the former muscles. Thus, two of the laryngeal muscles that would seem to be most important for control of voice F_0 with the PSR are the CT and TA. Therefore, one objective of the present study was to determine if the CT and TA muscles are also involved in the corrective vocal responses to pitch-shifted voice feedback.

With regard to a possible role of the CT and TA muscles in the PSR, it is important to know how rapidly they respond to pitch-shift stimuli. For these muscles to be involved in the PSR as measured at the peak of the response, it is important that they respond to the stimuli soon enough to be able to cause contraction of the muscles leading to the voice F_0 responses. Therefore, a secondary objective of this study was to measure the timing of the EMG responses to the pitch-shift stimuli. As a further test of the relationship between the EMG and voice F_0 responses, we tested the hypothesis that increases or decreases in EMG magnitudes corresponded with increases or decreases in magnitude of the voice F_0 responses. Results of the present study demonstrate that the CT and TA muscles are involved in the PSR, and the timing of their responses to pitch-shift stimuli indicates that they are fast enough to contribute to the observed voice F_0 responses.

II. METHODS

A. Subjects

Four healthy adults (age 22–60; one woman and three men) served as subjects. The three male subjects (HL, RB, and CRL) were authors of the study, who were familiar with the experimental protocol. None reported significant hearing loss, speech, language, or neurological disorder. This study was approved by the Northwestern University Institutional Review Board Human Subjects Committee.

B. Apparatus

Hooked-wire EMG electrodes were constructed by passing bipolar, insulated, stainless steel wires (.002 in. diameter) through the lumen of a 27 gauge hypodermic needle. Electrodes were steam sterilized prior to use. EMG potentials were amplified with Grass P511 amplifiers (10,000 gain; band-pass filtered at 60–3000 Hz).

Subjects were seated in a medical examination chair with AKG boom-set headphones and attached microphone (model K 270 H/C) placed on the head. The microphone signal was amplified with a MOTU Ultralite Mk3 firewire audio interface and processed for pitch shifting through an Eventide Eclipse Harmonizer. The pitch-shifted signal was amplified to a gain of 10 dB [sound pressure level (SPL)] greater than voice amplitude and fed back to the subject over the headphones. Acoustic calibrations were made with a Brüel & Kjær sound level meter (model 2250) and in-ear microphones (model 4100). A laboratory computer running MIDI software (Max/MSP v.4.1 by Cycling 74) was used to control the parameters of the pitch-shift stimulus, such as magnitude, duration, onset time, and interstimulus interval (ISI) through the Eventide Eclipse harmonizer. Microphone, headphone, transistor–transistor logic (TTL) control signals, and EMG potentials were low-pass filtered at 5 kHz, digitized (12 bit) at 10 kHz through a PowerLab A/D converter (model ML880, AD Instruments) and recorded on a second computer with Chart software (AD Instruments).

C. Procedures

Prior to insertion of electrodes, a subcutaneous injection of 0.5 ml xylocaine was made over the median cricothyroid ligament. Needle electrode assemblies were then percutaneously inserted into the CT and TA muscles. The needle was then withdrawn, leaving the hooked end of the wires within the muscles. The ends of the wires extending from the neck were connected to the amplifiers to verify correct placement by observing the EMG potentials on a computer screen. Accurate position in the desired muscle was verified with standard criteria (Hirano and Ohala, 1969; Hirano *et al.*, 1970). Following verification procedures, if the electrode did not appear to be in the correct muscle, the wires were withdrawn and new electrodes were inserted.

After verification of correct electrode placement, subjects were then instructed to sustain a vowel sound, /a/, at each of three different voice F_0 levels on different blocks of trials: Conversational pitch, a high pitch, but still within the head register, and the falsetto register. These general voice

F_0 levels were chosen because they encompass the range that has been studied in most previous studies, and because the voice F_0 responses to pitch-shifted voice feedback have been reported to increase in magnitude with increases in F_0 level (Liu and Larson, 2007). Also, pilot testing showed more clearly defined EMG responses at higher F_0 levels compared to the conversational level. At each voice F_0 level, subjects vocalized for approximately 6 s, ten times. During each vocalization, five pitch-shift stimuli were presented over the headphones. The first pulse in the sequence of five occurred between 500 and 1000 ms after vocal onset, and the succeeding pulses had an ISI varying between 700 and 900 ms. In each block of ten trials the pitch-shift stimulus was held constant in magnitude at either ± 100 or ± 300 cents (100 cents = 1 semitone), 200 ms duration. The sequencing of upward and downward stimuli was randomized within a block of trials. Over the course of the ten trials for each voice F_0 level, approximately 25 upward and 25 downward stimuli were presented. The order of the intervening three vocal tasks was randomized across all subjects.

Data were analyzed offline using IGOR PRO (v.5.0, Wavemetrics Inc.). First, the vocal and auditory feedback waveforms were converted to analog contours of voice F_0 using an autocorrelation method in Praat (Boersma, 2001). These signals were then converted to the cents scale using the formula: cents $= 100[39.86 \log 10 (f2/f1)]$, where $f1$ equals an arbitrary reference note at 195.997 Hz (G4) and $f2$ equals the voice F_0 in Hertz. The EMG signals were converted to a root-mean-square (RMS) measure with a window length of 25 ms and then digitally low-pass filtered at 20 Hz. Zero phase-shift filtering was applied by filtering in the forward and backward directions.

Ensemble signal averaging of the voice F_0, voice feedback and EMG waves was done by first segmenting each wave into epochs ranging from 200 ms before and 700 ms following the pitch-shift stimulus. Then, epochs in which the voice F_0 contours followed the stimulus direction (i.e., the same direction as the stimulus) and those that opposed the stimulus direction (i.e., opposite in direction to the stimulus) were tallied. For these two response categories, the category with the most number of epochs was chosen for signal averaging, which was triggered by a TTL pulse aligned with the pitch-shift stimulus onset. In most cases the predominant response direction of the voice F_0 contour in the epochs opposed the stimulus direction. The averaging algorithm then generated averages for the voice F_0, feedback and EMG RMS waves based on the predominant direction of the individual voice F_0 trials. Averaged voice F_0 contours that changed in the opposite direction to the stimulus were defined as opposing responses. Averages that changed in the same direction as the stimulus were defined as "following" responses. The same procedure was used to define EMG responses as opposing or "following." The reason for averaging only the epochs changing in a single direction was to generate averages that best represented the typical response mode rather than averaging different directions of responses. Averaging epochs in which the F_0 contour changed in different directions could potentially lead to an averaged response that registered no change.

From the average waveforms, a computer program measured the magnitude and latency of the voice F_0 and EMG responses at the time when the signals reached a maximum (upward changes) or minimum (downward changes) value. The criteria for acceptable F_0 and EMG average responses was that the peak had to occur at least 20 ms following the stimulus onset, that it was greater than (for upward responses) or less than (for downward responses) a value equal to twice the standard deviation of the prestimulus mean, and that the duration of the response that exceeded this value had to be at least 30 ms. In order to normalize the EMG data and thus compare the magnitude of EMG responses across muscles and subjects, EMG response percent magnitudes were calculated by dividing the peak magnitude of each EMG response by the mean of all responses for that specific muscle in all experimental conditions (i.e., voice F_0 level and stimulus magnitude). As a test of the hypothesis that the EMG responses contributed toward the voice F_0 responses, the peak times of the EMG responses were subtracted from the peak times of the voice F_0 responses (voice–EMG interval). Positive values indicated the EMG responses preceded the voice responses, and negative values indicated they followed the F_0 responses. Three-way ($3 \times 2 \times 2$) factorial analysis of variances (ANOVAs) were used to test the main effect of voice F_0 level (conversational, high and falsetto), stimulus direction (up and down) and magnitude (100 and 300 cents) along with their interactions on the magnitudes and latencies of EMG and vocal, and the voice–EMG interval. As the primary interest in this study was the latency and magnitude of the responses, vocal and EMG response statistics were done on the combined values of the opposing and following directions.

III. RESULTS

Across the three male subjects, the mean F_0 for the conversational level was 130 Hz, for the high level, 200 Hz, and for falsetto, 360 Hz. From the four subjects, recordings were made from six TA and six CT muscles. The falsetto frequency for the female subject, the only frequency at which we were able to get good averaged EMG responses, was 588 Hz. In most of the higher voice F_0 conditions (high and falsetto), there was a clear change in the EMG responses that occurred just prior to or during the vocal response. With the conversational voice condition, there were relatively few clear EMG responses to the pitch-shift stimulus. Across all subjects and conditions, 68 EMG responses were measured, with 31 in the falsetto, 20 in the high, and 17 in the conversational voice condition. Figure 1 shows representative EMG records of the left CT and left TA from one of the subjects while performing a vocal glissando. The activity of each muscle increases as the voice F_0 increases, and they both decrease as F_0 decreases. These changes in CT and TA EMG potentials associated with changes in voice F_0 are typical of those reported in previous studies (Hirano and Ohala, 1969; Hirano et al., 1970).

Figures 2 and 3 illustrate averaged records from one subject in the three voicing conditions with 100-cent stimuli. Figures 4 and 5 show results for the 300-cent stimuli. In
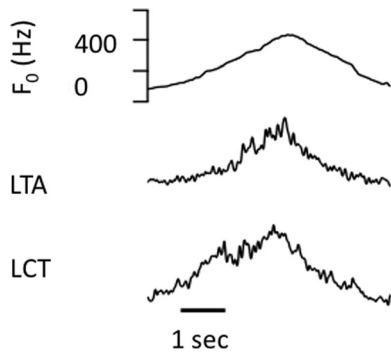
FIG. 1. Traces representing voice F_0 contour, top, RMS voltage of left thyroarytenoid EMG (second trace) and left cricothyroid EMG (third trace) from one of the subjects performing an upward and downward vocal glissando.

each of Figs. 2–5, with the downward stimulus, there is a compensatory increase in voice F_0, and with the upward stimulus, a decrease in F_0. Figures 2–5 also show conspicuous EMG responses in the falsetto voice condition, which changed in the opposite direction to the stimulus. EMG responses in the high and conversational voice F_0 levels were quite small in comparison.

Figure 6 shows box plots of the voice–EMG intervals for stimulus direction, magnitude, and voice F_0 level. A three-way ANOVA performed on the voice–EMG interval for the three conditions yielded nonsignificance for stimulus direction [$F(1,56) = 2.87$, $p = 0.095$] and stimulus magnitude [$F(1,56) = 1.54$, $p = 0.219$], but a significant effect for the voice F_0 condition [$F(2,56) = 5.04$, $p = 0.022$]. Post-hoc Bonferroni tests showed that in the falsetto condition, voice–EMG intervals (mean 90 ms) were significantly longer than the conversational condition (mean −58 ms; $p = 0.026$). The difference in these means shows that in the falsetto condition, the EMG responses occurred prior to the vocal responses, whereas in the conversational condition, the EMG

responses reached a peak following the voice responses. The mean of the voice–EMG interval for the high voice condition was 18 ms, indicating that the EMG responses also preceded the voice response. There were no significant interactions across the three independent variables.

Figure 7 shows box plots of vocal response magnitude and latency as a function of stimulus direction, magnitude, and voice conditions. A three-way ANOVA performed on vocal response magnitudes against the voice F_0 condition [$F(2, 23) = 0.411$, $p = 0.668$], stimulus direction: $F(1, 23) = 0.630$, $p = 0.436$) and stimulus magnitude $F(1, 23) = 0.151$, $p = 0.701$) failed to show any significant differences. Likewise, a three-way ANOVA performed on vocal response latency against the voice F_0 condition [$F(2, 23) = 1.098$, $p = 0.350$], stimulus direction [$F(1, 23) = 0.445$, $p = 0.511$], and stimulus magnitude [$F(1, 23) = 0.025$, $p = 0.875$] failed to show any significant differences. For both response magnitude and latency, no significant interactions among three variables were found ($p > 0.05$).

Figure 8 shows box plots of the percent magnitude of the EMG response and latency as a function of stimulus direction, stimulus magnitude, and voice F_0 condition. A three-way ANOVA performed on EMG response percent magnitudes showed a significant increase with the voice F_0 condition [$F(2, 56) = 5.560$, $p = 0.006$], where the falsetto level yielded significantly larger percent muscle responses than the conversational level ($p = 0.005$, Bonferroni correction). EMG percent magnitudes did not differ significantly as a function of stimulus direction [$F(1, 56) = 0.897$, $p = 0.348$] or magnitude [$F(1, 56) = 0.376$, $p = 0.542$]. There was a significant interaction for EMG percent magnitudes between stimulus direction and voice F_0 condition [$F(2, 56) = 3.240$, $p = 0.047$], but there were no significant interactions for the other conditions. Due to the significant interaction between stimulus direction and voice F_0 condition, separate two-way ANOVAs (stimulus magnitude and voice condition) were performed on
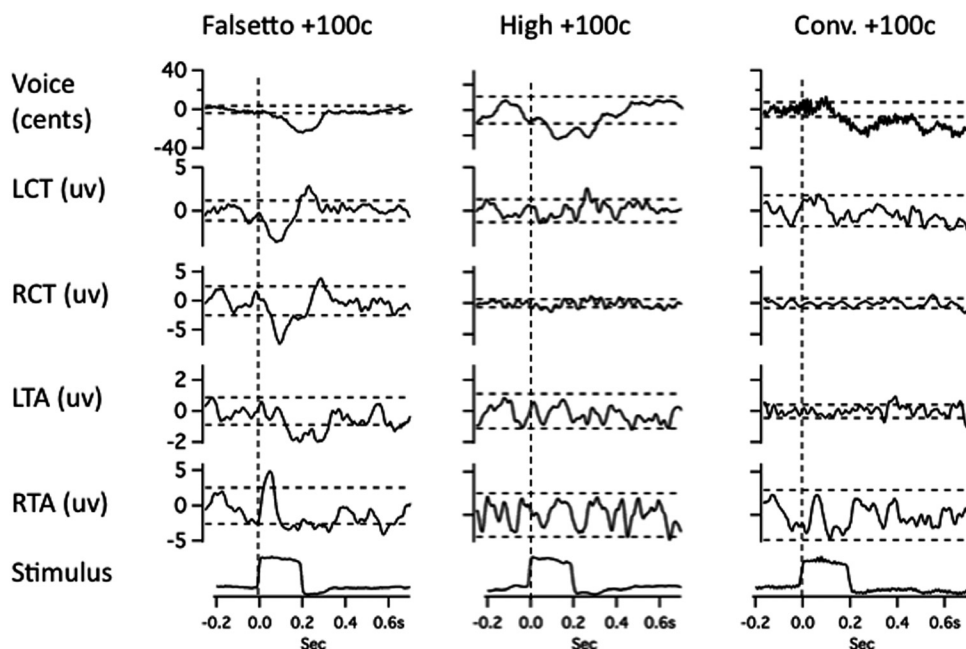

FIG. 2. Averaged traces of voice F_0 (top trace), RMS voltage of left cricothyroid, right cricothyroid, left thyroarytenoid, right thyroarytenoid, and the stimulus artifact trace (bottom) following 100-cent upward pitch-shift stimulus. Left column shows responses during falsetto production, middle column high F_0, and right column conversational F_0 production. Horizontal dashed lines represent ±2 standard deviations of the prestimulus mean amplitude. Vertical dashed lines show stimulus onset time.
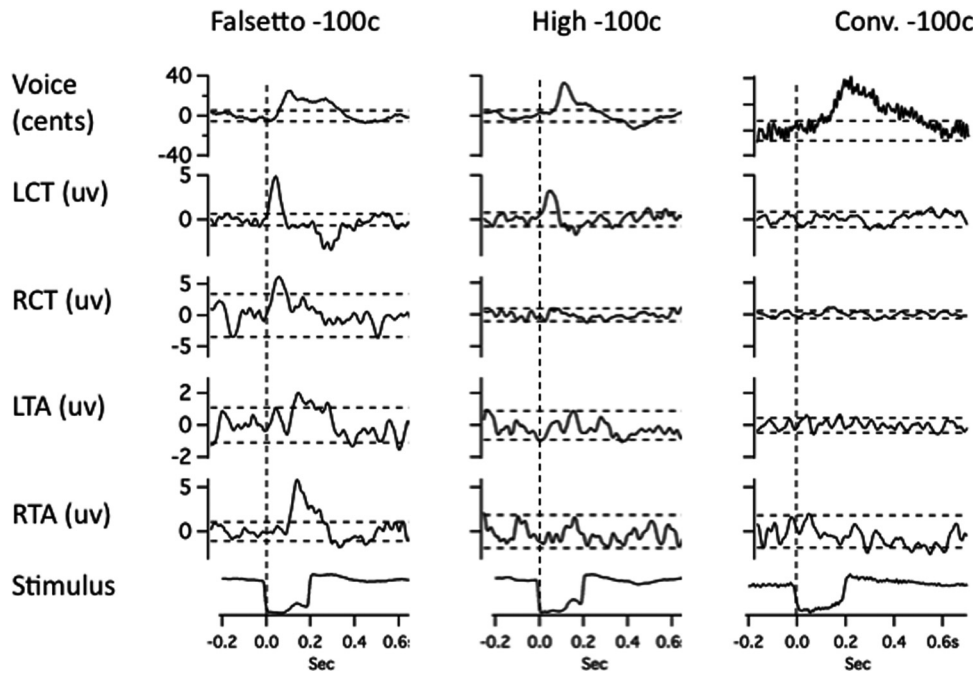
FIG. 3. Averaged traces of voice F_0 (top trace), RMS voltage of left cricothyroid, right cricothyroid, left thyroarytenoid, right thyroarytenoid, and the stimulus artifact trace (bottom) following 100-cent downward pitch-shift stimulus. Left column shows responses during falsetto production, middle column high F_0, and right column conversational F_0 production. Horizontal dashed lines represent ±2 standard deviations of the prestimulus mean amplitude. Vertical dashed lines show stimulus onset time.

the EMG percent magnitude for the upward and downward conditions. The results for the upward stimulus direction showed the stimulus magnitude to be not significant $[F(1, 28) = 0.012, p = 0.915]$, but voice F_0 condition was significant $[F(2, 28) = 8.892, p = 0.001]$, where Bonferroni tests showed significantly larger percent muscle responses for the falsetto level than both the conversational $(p = 0.002)$ and high levels $(p = 0.027)$. For the downward stimulus directions, stimulus magnitude $[F(1, 28) = 0.545, p = 0.466]$, voice F_0 condition $[F(2, 28) = 1.006, p = 0.379]$, and stimulus magnitude by voice condition $[F(2, 28) = 0.291, p = 0.750]$ were not significant.

Figures 2–5 also show that most voice and EMG responses changed in the opposite direction to the stimulus.

Traditionally, it has been found that most voice responses are in the opposite direction to the stimuli and are termed "opposing" responses. Those that change in the same direction, a minority, are termed "following" responses (Burnett *et al.*, 1998). Table I shows that all of the responses in the falsetto voice level changed in the opposite direction to the stimuli along with most of the responses in the high level. The EMG responses in the conversational voice level were mixed for the upward and downward stimuli. Table II shows the number and percentage of EMG responses that changed in the same versus the opposite direction as the vocal response. In the falsetto voice, all of the EMG responses were in the same direction as the voice responses, whereas for the conversational and high voice F_0 levels there was a
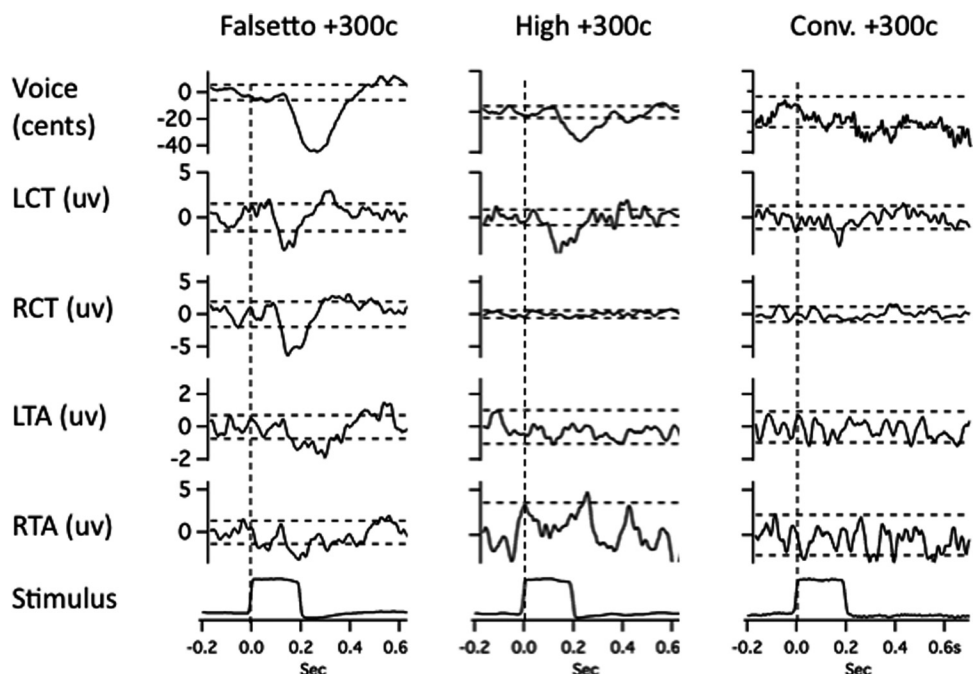

FIG. 4. Averaged traces of voice F_0 (top trace), RMS voltage of left cricothyroid, right cricothyroid, left thyroarytenoid, right thyroarytenoid, and the stimulus artifact trace bottom following 300-cent upward pitch-shift stimulus. Left column shows responses during falsetto production, middle column high F_0, and right column conversational F_0 production. Horizontal dashed lines represent ±2 standard deviations of the prestimulus mean amplitude. Vertical dashed lines show stimulus onset time.
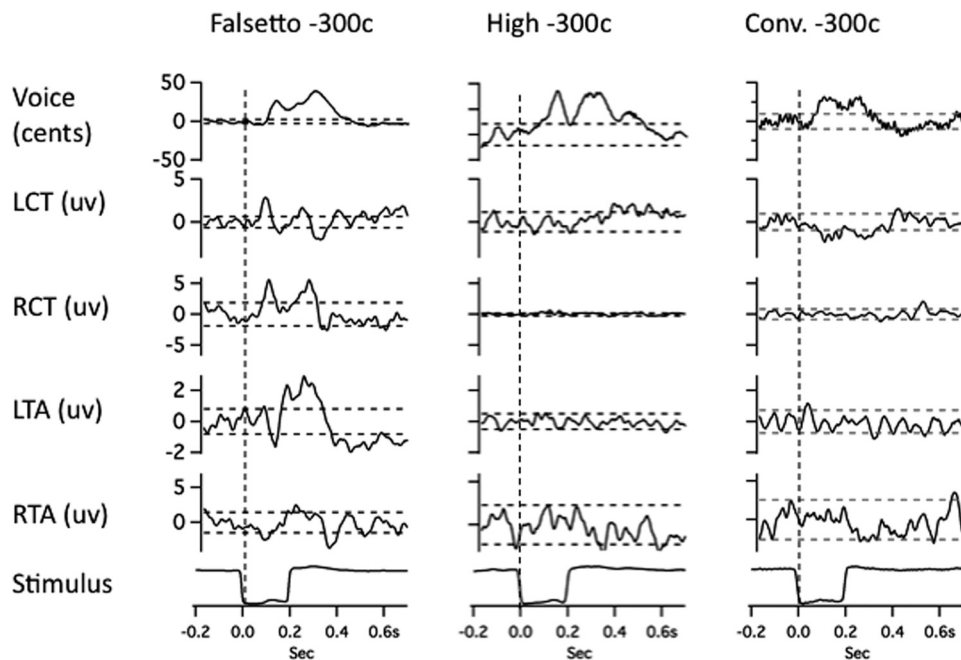
Falsetto -300c   High -300c   Conv. -300c

FIG. 5. Averaged traces of voice F_0 (top trace), RMS voltage of left cricothyroid, right cricothyroid, left thyroarytenoid, right thyroarytenoid, and the stimulus artifact trace bottom following 300-cent downward pitch-shift stimulus. Left column shows responses during falsetto production, middle column high F_0, and right column conversational F_0 production. Horizontal dashed lines represent ± 2 standard deviations of the prestimulus mean amplitude. Vertical dashed lines show stimulus onset time.

35/65% and a 70/30% ratio of same direction to opposite, respectively. Thus, the EMG responses in the falsetto level changed in the opposite direction to the stimulus and the same direction as the voice response, and hence they showed the clearest relationship with the voice F_0 responses.

The close relationship between the EMG and F_0 averages was sometimes violated in this study. Figure 9 shows data for one subject in which there is a slight decrease in voice F_0 followed by a relatively larger increase in F_0 (downward stimulus, right). In this case, the left CT shows an increase in activity almost simultaneously with a decrease in activity in the right CT muscle. For an increase in voice pitch feedback (left), the opposite trend in EMG responses occurs. Thus, in both cases, the left and right CT muscles change their activity opposite to one another. Another example in Fig. 10 shows data from one subject in which the averaging program was set to accept *all* responses for averaging (All; far right), just the *upward* F_0 responses, (Up, center), or just the *downward* F_0 responses (down, left) to the downward pitch-shift stimuli. The result is that by accepting indi-

vidual trials that opposed the stimulus direction (up, $N = 13$), the F_0 response is larger than when all responses were accepted for the average ($N = 25$). When just the down trials are accepted ($N = 12$), the F_0 response goes down and follows the stimulus direction. Despite the variation in the F_0 responses, the EMG responses, especially those from the left TA are remarkably consistent across the three averages. The observations shown in Figs. 9 and 10 reveal considerable complexity in the relation between laryngeal EMG recordings and changes in voice F_0 following pitch-shift stimuli. There does not appear to be a simple direct relation between any one muscle and its effect on the voice F_0 responses.

IV. DISCUSSION

The present study was designed to test the hypothesis that the CT and TA laryngeal muscles are involved in voice F_0 responses to pitch-shifted voice auditory feedback. Further, we tested the hypothesis that increases or decreases in the activity of these muscles corresponded to similar changes
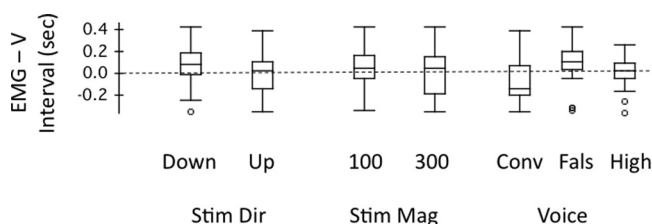

FIG. 6. Box plots of voice–EMG interval (s) plotted against downward and upward stimulus directions, 100- and 300-cent stimulus magnitudes, and the conversational, falsetto and high voice F_0 levels. Horizontal dashed line indicates division between intervals where EMG preceded the vocal response top, and followed the vocal response bottom. Box plot definitions: middle line is median, top and bottom of boxes are 75th and 25th percentiles, whiskers extend to limits of main body of data defined as high hinge +1.5 (high hinge – low hinge), and low hinge –1.5 (high hinge – low hinge); outliers are indicated by small "○."
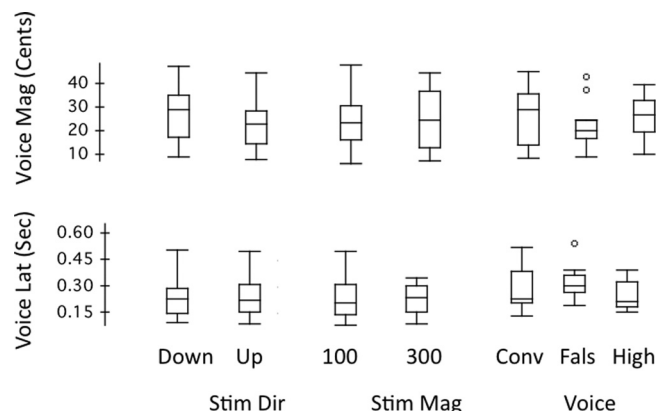

FIG. 7. Box plots of voice response magnitude (cents) and voice response peak latency (s) for stimulus direction, stimulus magnitude, and voicing condition. See Fig. 6 for box plot definitions.
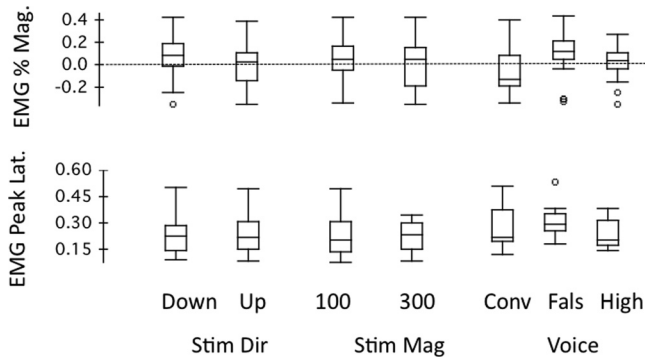
FIG. 8. Box plots of percent EMG response magnitude and EMG peak latency against stimulus direction, stimulus magnitude, and voice conditions. See Fig. 6 for box-plot definitions.

in voice F_0. Numerous studies have previously shown that these muscles are used for voluntary control of voice F_0, and it has been shown that these muscles are excited reflexively by sensory stimulation of laryngeal nerves or tissues (Kirchner and Suzuki, 1968; Suzuki and Sasaki, 1977; Ludlow et al., 1992; Ludlow et al., 1995; Barkmeier et al., 2000; Andreatta et al., 2002). Results of the present study have shown these muscles are also involved in reflexive vocal responses to pitch-shifted voice feedback for voicing in the falsetto register. Specifically, an upward shift in pitch feedback caused a reduction in voice F_0, CT and TA EMG, whereas a decrease in voice pitch feedback caused an increase in voice F_0, CT, and TA EMG activity for the falsetto voice.

The timing and directional sensitivity of the voice and EMG responses suggests they are reflexive. As shown in Fig. 6, in the falsetto register, most EMG responses preceded the vocal response. However, it is difficult to compare the vocal peak latencies of this study with previous studies as most previous studies did not report peak latencies. Kiran and Larson (2001) reported peak times of 256 ms for voice F_0 responses to pitch-shifted voice feedback, which are similar to results of the present study. It might be inferred therefore, that the latencies of response onset times in the present study were similar to those reported previously, or 100–150 ms (Burnett et al., 1998; Hain et al., 2000; Natke et al., 2003; Bauer et al., 2006; Chen et al., 2007). Vocal response latencies such as these are most likely of an automatic nature because they are directionally sensitive and too fast to be voluntarily generated. For example, response latencies in a choice voluntary reaction-time test are usually in the range of 300 ms or more (Rosenbaum, 1991). Simple reaction

TABLE I. Counts of opposing (OPP) and following (FOL) voice and EMG responses with respect to stimulus direction across voice F_0 conditions.

		Conversational	Falsetto	High	Total
Voice[a]	FOL	1	0	1	2
	OPP	9	11	13	33
EMG[b]	FOL	9	0	5	14
	OPP	8	31	15	54

[a] $\chi^2 = 1.061$, df $= 2$, $p = 0.588$.
[b] $\chi^2 = 19.16$, df $= 2$, $p = 0.0001$.

TABLE II. Number (%) of EMG response directions that changed in the same or opposite direction as the vocal response.

	EMG same direction	EMG opposite direction
Conversational	6 (35%)	11 (65%)
High	14 (70%)	7 (30%)
Falsetto	31 (100%)	0

times requiring no choice can be as fast as 100 ms (Luschei et al., 1967). As the voice F_0 responses in the pitch-shift paradigm are similar to a "choice" task, being directionally sensitive, they are probably too fast to be produced voluntarily. It then follows that the EMG responses that occur prior to the vocal responses are also of an automatic nature. This is not to say that the responses cannot be influenced by voluntary mechanisms (see Xu et al., 2004; Chen et al., 2007; Liu et al., 2007), but volitional variables were not a factor in the present study.

Another factor supporting the reflex hypothesis is that the EMG responses either increased or decreased along with the corresponding F_0 responses, at least for the falsetto voice condition (see Figs. 2–5 and Table I). Moreover, Table I shows that for the falsetto voice, all EMG responses opposed the stimulus direction. Although most of the EMG responses in the high condition also opposed the stimulus directions, the responses in the conversational condition were split between the stimulus directions. Figure 6 also shows that in the falsetto condition, the voice–EMG intervals were positive, indicating that the EMG responses occurred before the vocal response. During the high condition, and especially the conversational condition, the EMG potentials did not always precede the vocal responses. Thus, all of the data from the falsetto condition show a very clear relationship between EMG responses, stimulus direction, and vocal responses.

Although the various measurements show a clear relationship between EMG and vocal responses for the falsetto condition, this is not the case for the conversational and high voice F_0 levels. Although we do not have a definitive
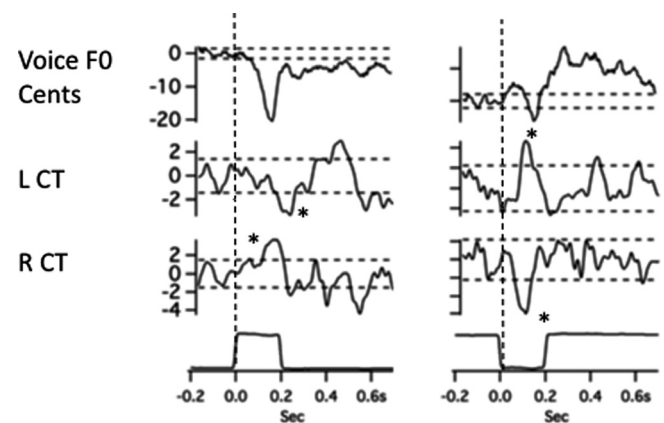

FIG. 9. Averaged traces of voice F_0 (top trace), RMS voltage of left cricothyroid, right cricothyroid, and the stimulus artifact trace (bottom) following 100-cent upward pitch-shift stimulus left, and downward stimulus, right. Asterisks denote peak and trough of responses described in text. Horizontal dashed lines represent ± 2 standard deviations of the prestimulus mean amplitude. Vertical dashed lines show stimulus onset time.
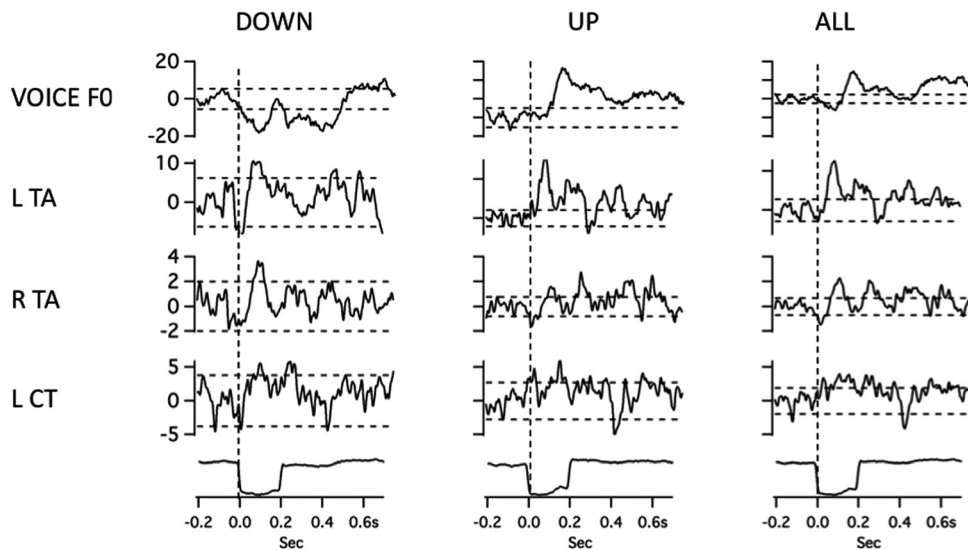
FIG. 10. Averaged traces of voice F_0 (top trace), RMS voltage of left thyroarytenoid, right thyroarytenoid, left cricothyroid, and the stimulus artifact trace (bottom), following a 100-cent downward pitch-shift stimulus. Left column shows responses obtained by selecting only downward F_0 traces for averaging, middle column for upward traces, and right column selecting both upward and downward responses for averaging. Horizontal dashed lines represent ± 2 standard deviations of the prestimulus mean amplitude. Vertical dashed lines show stimulus onset time.

explanation of why the EMG responses in the conversational and high F_0 levels were poorly related to the vocal responses, it is reasonable to speculate that they represent the normal variation in muscle activity that is observed in most muscles. Moreover, as it is known that some motor units in the CT muscle are inactive at a conversational voice F_0 level, activate at higher F_0 levels and continue to increase their discharge rate at higher frequencies (Sutton *et al.*, 1972), it is reasonable to suppose that there was very little muscle activity present at the conversational levels, and the small electrodes used in this study failed to record muscle activations at the lower pitch levels.

A related question is why in the falsetto voice F_0 level the EMG potentials should have been so prominent compared to the conversational F_0 level, especially considering that the magnitudes of the vocal responses were similar across all the voice conditions? First, it can be said that in general there was a much greater level of EMG activity in the muscles with the higher voice F_0 levels. Note the magnitude of the EMG levels in the falsetto condition in Figs. 2–5 compared with the high and conversational F_0 levels. Previous studies have also shown that with voluntary increases in voice F_0 level, there is an increase in the amplitude of EMG signals (Hirano and Ohala, 1969; Hirano *et al.*, 1970; Sutton *et al.*, 1972). In order to vocalize at a high F_0, laryngeal muscles must contract more forcefully to stretch (CT) and increase the stiffness (CT and TA) of the vocal folds. Vocalizing at low F_0 levels, which requires less forceful muscle contraction, means that a relatively small set of small motor units of a muscle would be activated with a relatively low discharge rate (Sutton *et al.*, 1972). Therefore, at low F_0 levels, a pitch-shift stimulus may only increase activation levels of a few active motor units, whereas other motor units would remain in an inactive state. At higher F_0 levels (e.g., falsetto), most if not all of the motor units may be activated, and a small excitatory input to the motor neuron pool resulting from a pitch-shift stimulus, might change the activation levels of all active motor units, which would have a more demonstrable effect on the EMG activity. It should also be remembered in the context of this discussion, that the magni-

tude of the vocal responses under all conditions studied was rather small, i.e., approximately 30 cents, which for a F_0 level of 130 Hz, is a change in 2–3 Hz. Therefore, it is to be expected that the overall change in muscle contraction and EMG levels would be rather small for all experimental conditions tested.

The observations shown in Figs. 9 and 10 reveal other properties of this reflexive system that, to our knowledge, have not been previously published. As shown in Fig. 9, one CT muscle showed a reduction in activity whereas the other CT showed an increase following a pitch-shift stimulus. These oppositely changing EMG patterns suggest that inputs to one of the muscles resulted from mechanisms involved in producing a compensatory voice response, whereas inputs to the other muscle were related to mechanisms producing a following response. In other words, these responses suggest that two contrary vocal control mechanisms were simultaneously active. The observations in Fig. 10 suggest that within a set of 25 pitch-shift stimuli, compensatory and following mechanisms for vocal control were switching back and forth. Taken together, these observations suggest an order of complexity in vocal control that requires additional study in order to understand neuromuscular control of the voice.

Although it has been speculated that following responses may be due to treating the feedback signal itself as the referent (Hain *et al.*, 2000) or that they may result from misperception of the stimulus (Larson *et al.*, 2007), neither of these two studies provide definitive causes of these responses. However, results of the present study suggest that mechanisms involved in the motor control of laryngeal muscles may be partly responsible for the following responses. That is, if the causes were related to "perception" of the stimulus, one would expect both muscles to exhibit identical patterns of contraction. On the other hand, if the mechanisms were purely motor (motor neurons or muscles), it is difficult to understand how the same stimulus could cause the two muscles to behave differently. A further possibility is that the neural mechanisms involved in the translation from stimulus perception to motor output may be involved in these responses. As we do not understand these neural mechanisms,

we cannot conjecture further on the causes of the anomalous EMG responses reported here. Further research is necessary in order to understand these EMG responses, as well as other details related to generating vocal motor responses to changes in voice auditory feedback.

V. CONCLUSION

This study has demonstrated that both the CT and TA laryngeal muscles are involved in the reflexive control of voice F_0. When vocalizing in the falsetto register, the muscles generally changed their activity in the same direction as that of voice F_0, meaning that they increased their activity for an increase in voice F_0 and decreased their activity with a decrease in F_0. The latency of the peak magnitude of the EMG responses occurred about 80 ms prior to the peak of the vocal response. Measures of EMG responses in relation to the vocal responses in the high or conversational voice F_0 conditions exhibited greater variation in the timing, magnitude, and direction of the EMG responses. Thus, at least for the falsetto register, the CT and TA muscles showed a pattern of activity consistent with a role of generating changes in voice F_0 in response to perturbations in voice pitch auditory feedback. Variations in patterns of EMG responses between muscles of a single subject provide insight into factors that may contribute to variability in F_0 response measures across different subjects.

ACKNOWLEDGMENTS

Andreatta, R. D. Mann, E. A. Poletto, C. J., and Ludlow, C. L. (**2002**). "Mucosal afferents mediate laryngeal adductor responses in the cat," J. Appl. Physiol. **93**, 1622–1629.

Barkmeier, J. M. Bielamowicz, S. Takededa, N., and Ludlow, C. L. (**2000**). "Modulation of laryngeal responses to superior laryngeal nerve stimulation by volitional swallwing in awake humans," J. Neurophysiol. **83**, 1264–1272.

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* (Wiley-Blackwell, Hoboken, NJ), pp. 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**, 3153–3161.

Chanaud, C., and Ludlow, C. (**1992**). "Single motor unit activity of human intrinsic laryngeal muscles during respiration," Ann. Oto. Rhinol. Laryngol. **101**, 832–840.

Chen, S. H. Liu, H. Xu, Y., and Larson, C. R. (**2007**). "Voice F0 responses to pitch-shifted voice feedback during English speech," J. Acoust. Soc. Am. **121**, 1157–1163.

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

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.

Hirano, M., and Ohala, J. (**1969**). "Use of hooked-wire electrodes for electromyography of the intrinsic laryngeal muscles," J. Speech Hear. Res. **12**, 362–373.

Hirano, M. Vennard, W., and Ohala, J. (**1970**). "Regulation of register, pitch and intensity of voice," Folia Phoniatr. Logo. **22**, 1–20.

Hirose, H., and Gay, T. (**1972**). "The activity of the intrinsic laryngeal muscles in voicing control," Phonetica **25**, 140–164.

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

Kawahara, H. (**1995**). "Hearing voice: Transformed auditory feedback effects on voice pitch control," in *Computational Auditory Scene Analysis and International Joint Conference on Artificial Intelligence*.

Kiran, S., and Larson, C. R. (**2001**). "Effect of duration of pitch-shifted feedback on vocal responses in Parkinson's Disease patients and normal controls," J. Speech. Lang. Hear. Res. **44**, 975–987.

Kirchner, J. A., and Suzuki, M. (**1968**). "Laryngeal reflexes and voice production," in *Annals of New York Academy of Sciences* (Annals of the New York Academy of Sciences, New York), pp. 98–109.

Larson, C. R. Kempster, G. B., and Kistler, M. K. (**1987**). "Changes in voice fundamental frequency following discharge of single motor units in cricothyroid and thyroarytenoid muscles," J. Speech. Hear. Res. **30**, 552–558.

Larson, C. R. Sun, J., and Hain, T. C. (**2007**). "Effects of simultaneous perturbations of voice pitch and loudness feedback on voice F0 and amplitude control," J. Acoust. Soc. Am. **121**, 2862–2872.

Lindestad, P.-Å., Fritzell, B., and Persson, A. (**1990**). "Evaluation of laryngeal muscle function by quantitative analysis of the EMG interference pattern," Acta Otolaryngol. (Stockh) **109**, 467–472.

Lindestad, P.-Å., Fritzell, B., and Persson, A. (**1991**). "Quantitative analysis of laryngeal EMG in normal subjects," Acta Otolaryngol. (Stockh) **111**, 1146–1152.

Liu, H., and Larson, C. R. (**2007**). "Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex," J. Acoust. Soc. Am. **122**, 3671–3677.

Liu, H. Zhang, Q. Xu, Y., and Larson, C. R. (**2007**). "Compensatory responses to loudness-shifted voice feedback during production of Mandarin speech," J. Acoust. Soc. Am. **122**, 2405–2412.

Loucks, T. M. Poletto, C. J. Saxon, K. G., and Ludlow, C. L. (**2005**). "Laryngeal muscle responses to mechanical displacement of the thyroid cartilage in humans," J. Appl. Physiol. **99**, 922–930.

Ludlow, C. Van Pelt, F., and Koda, J. (**1992**). "Characteristics of late responses to superior laryngeal nerve stimulation in humans," Ann. Otol. Rhinol. Laryngol. **101**, 127–134.

Ludlow, C. L. and Lou, G. (**1996**). "Observations on human laryngeal muscle control," in *Vocal Fold Physiology: Conrolling Complexity and Chaos*, edited by P. J. Davis, and N. H. Fletcher (Singular, San Diego), pp. 201–218.

Ludlow, C. L. Schulz, G. M. Yamashita, T., and Deleyiannis, F. W.-B. (**1995**). "Abnormalities in long latency responses to superior laryngeal nerve stimulation in adductor spasmodic dysphonia," Ann. Oto. Rhinol. Laryngol. **104**, 928–935.

Luschei, E. L.O., R., Finnegan, E. M. Baker, K. K., and Smith, M. E. (**2006a**). "Patterns of laryngeal electromyography and the activity of the respiratory system during spontaneous laughter," J. Neurophysiol. **96**, 442–450.

Luschei, E. Saslow, C., and Glickstein, M. (**1967**). "Muscle potentials in reaction time," Exp. Neurol. **18**, 429–442.

Luschei, E. S. Ramig, L. O. Baker, K. L., and Smith, M. E. (**1999**). "Discharge characteristics of laryngeal single motor units during phonation in young and older adults and in persons with parkinson disease," J. Neurophysiol. **81**, 2131–2139.

Luschei, E. S. Ramig, L. O. Finnegan, E. M. Baker, K. K., and Smith, M. E. (**2006b**). "Patterns of laryngeal electromyography and the activity of the respiratory system during spontaneous laughter," J. Neurophysiol. **96**, 442–450.

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.

Rosenbaum, D. A. (**1991**). *Human Motor Control* (Academic, New York).

Sataloff, R. T. Mandel, S. Heman-Ackah, Y., and Manon-Espaillat, R. (**2006**). *Laryngeal Electromyogrphy* (Plural, San Diego).

Sutton, D. Larson, C. R., and Farrell, D. M. (**1972**). "Cricothyroid motor units," Acta. Otolaryngol. **74**, 145–151.

Suzuki, M., and Sasaki, C. (**1977**). "Effect of various sensory stimuli on reflex laryngeal adduction," Ann. Oto. Rhinol. Laryngol. **86**, 30–36.

Titze, I. R. (**1994**). *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ).

Xu, Y. Larson, C. Bauer, J., and Hain, T. (**2004**). "Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences," J. Acoust. Soc. Am. **116**, 1168–1178.