Age-related differences in vocal responses to pitch feedback perturbations: A preliminary study

Hanjun Liu

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

Nicole M. Russo

Department of Pediatrics, Department of Behavioral Sciences, and RUSH Neurobehavioral Center, RUSH University Medical Center, 4711 West Golf Road, Skokie, Illinois 60076

Charles R. Larson^{a)}

Department of Communication Science and Disorders, Northwestern University, 2240 Campus Drive, Evanston, Illinois 60208

(Received 24 June 2009; revised 12 November 2009; accepted 14 November 2009)

The present study tested the effect of age on voice fundamental frequency (F_0) responses to pitch-shifted feedback. Pitch-shift stimuli (−100 cents, 200 ms duration) were presented to 30 native-English speakers: 10 children (7–12 yrs), 10 younger adults (19–21 yrs), and 10 older adults (60–73 yrs). Significantly larger response magnitudes were found in the older group compared to the children and young adult groups, while the longest latencies were associated with the children group as compared to the two adult groups. These findings provide preliminary evidence of an age effect on the modulation of vocal responses to pitch-shifted feedback.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3273880]

PACS number(s): 43.70.Ep, 43.71.Ft, 43.71.Qr, 43.71.Rt [AL]

Pages: 1042–1046

I. INTRODUCTION

The importance of auditory feedback during speech production has been well known (Lombard, 1911; Lane and Tranel, 1971) for many years. In recent years, the frequency perturbation technique was developed and applied to the investigation of the mechanisms underlying auditory feedback control of voice fundamental frequency (F_0) (Elman, 1981; Kawahara, 1994; Burnett *et al.*, 1997). In this protocol, voice pitch auditory feedback is unexpectedly shifted upward or downward and presented to the speakers during voice or speech production. Numerous studies have demonstrated that subjects compensate for these perturbations in pitch feedback by changing their voice F_0 in the opposite direction to the stimulus during ongoing production of speech and voice (Burnett *et al.*, 1998; Donath *et al.*, 2002; Natke *et al.*, 2003; Xu *et al.*, 2004; Chen *et al.*, 2007; Liu and Larson, 2007; Liu *et al.*, 2009). These studies suggest that auditory feedback can be used to stabilize voice F_0 around a desired level and facilitate accurate and timely adjustment in voice control during speech.

The task-dependent activity of the audio-vocal system has been demonstrated by observing the within-subject effect on the modulation of vocal F_0 responses such as vocal parameters (voice F_0 level) (Liu and Larson, 2007), auditory stimuli (duration, velocity, etc.) (Burnett *et al.*, 1998; Larson *et al.*, 2000), and experimental tasks (singing, speaking, etc.) (Natke *et al.*, 2003; Xu *et al.*, 2004; Chen *et al.*, 2007). However, only a few studies investigated the auditory feed-

back processing of voice F_0 from the perspective of different subject groups. For example, Kiran and Larson (2001) showed that the vocal responses to pitch-shifted feedback in individuals with Parkinson's disease had significantly longer peak times and end times than those of control subjects. It was reported that some children with autism spectrum disorders (ASDs) showed abnormal vocal responses to pitch perturbations as compared with typically developing children development (Russo *et al.*, 2008). During the exposure to the brief pitch perturbations in auditory feedback, singers compensated to a lesser degree than nonsingers (Jones and Kough, 2008). Moreover, singers produced significantly higher voice F_0 values during the initial test-trials that occurred after pitch-shift trials than those produced during baseline and control trials, which was not observed for nonsingers. These studies suggest that between-subject effects may contribute to the modulation of vocal F_0 responses.

One important remaining question that has not been previously addressed is the effect of age on the auditory feedback control of voice F_0 . Age-related changes, as is well known, impact both structure and function of the voice and speech mechanisms. The effects of aging on speech production have been extensively investigated to identify the specific acoustic changes that occur in the speech of aging adults (Benjamin, 1981; Ramig and Ringel, 1983; Benjamin, 1997; Mueller, 1997; Sataloff *et al.*, 1997). It has been noted that, as compared to younger adults, acoustic characteristics associated with advanced age include slower speaking and reading rates, higher standard deviation for voice F_0 , greater jitter and spectral noise, and lower vowel formants (Shipp and Hollien, 1969; Wilcox and Horii, 1980; Ramig and Ringel, 1983). These changes, generally, are attributed to the

^{a)}Author to whom correspondence should be addressed. Electronic mail: clarson@northwestern.edu

physiological changes in speech mechanisms, reduced sensory feedback, decreased speed/accuracy of motor control, and diminished cognitive-linguistic function (Liss *et al.*, 1990; Torre and Barlow, 2009).

Schneider *et al.* (2009) noted that “many essential characteristics of vocal habits and several risk factors for voice disorders later in the professional life are already established in childhood.” Special care has been given to the pediatric studies of voice-related disorders, but there have been fewer studies of the normal childhood voice. It is generally agreed that voice F_0 in normal children decreases with age (Bennett, 1983; Glaze *et al.*, 1988, 1990; Nicollas *et al.*, 2008). Conflicting findings, however, in measures of jitter and shimmer were reported. Nicollas *et al.* (2008) found no significant age effect on jitter and shimmer in normal children, which is in contrast with the findings reported by Glaze and colleagues (1988, 1990).

With respect to previous studies using the pitch-shift feedback paradigm, most studies were either done with younger adults or age-matched subjects. How other populations such as normal children respond to pitch-shifted voice feedback in comparison to older adults is unknown. Although all of these groups produce compensatory responses to voice pitch feedback perturbations like the younger adults (Kiran and Larson, 2001; Russo *et al.*, 2008), no between-subject comparisons have been performed to see if there are age-related differences in auditory feedback control of voice F_0 . This comparison is important because it will provide information on voice development before laryngeal maturation and the aging-specific physiological changes that affect the processing of auditory feedback during ongoing vocalization. Ultimately, this information will help us understand the complex mechanisms underlying auditory feedback control of voice F_0 .

The primary goal of this study was to compare vocal responses to voice pitch-shifted feedback between three groups of native-English speakers, children, young adults, and older adults. The data set for this study came from three previous studies conducted in our laboratories over recent years (Chen *et al.*, 2007; Liu *et al.*, 2008; Russo *et al.*, 2008). Because the methodology differed slightly between each study, we focused on one specific experimental paradigm that was common to all three studies. We tested the hypothesis that age would affect the latency and/or the magnitude of voice F_0 responses to pitch-shifted voice feedback.

II. METHOD

A. Subjects

Three age groups of native-English speaking subjects participated in the study: children ($N=10$, 7–12 years old), young adults ($N=10$, 19–21 years old), and older adults ($N=10$, 60–73 years old). Because this study was not restricted by gender and there were no known effects of gender on the pitch-shift reflexes, these subjects were not gender-matched. None of the children or the younger adults reported a history of speech, hearing, or neurological disorders. All of the older adults passed the hearing screening at 250, 500, 1000, and 2000 Hz at 25 dB hearing level (HL); at 4000 Hz, seven

passed the screening at 25 dB HL and three at 50 dB HL. Hearing screening was also performed on children, and they all passed at the threshold of 20 dB HL for pure tone frequencies of 250–8000 Hz. All of the subjects signed informed consent approved by the Northwestern University Institutional Review Board. Additionally, all adults and both the parents of the children and the children themselves signed consent and assent for participation in the study.

B. Apparatus

The younger and the older adults were tested in a sound-treated room (IAC booth, model 1201), where they wore Sennheiser headphones with attached microphone (model HMD 280) throughout the testing. The children were tested in a laboratory room. Although the ambient noise level in the latter two rooms was not as tightly controlled as the sound treatment room, it was previously shown that the relative loudness of voice feedback or the addition of varying levels of masking noise had no significant effect on the amplitude or latencies of responses to pitch-shifted feedback (Burnett *et al.*, 1998). Before the testing, a Brüel and Kjaer sound level meter (model 2250) and in-ear microphones (model 4100) were used to adjust the acoustical feedback pathway to a gain of 10 dB sound pressure level between the voice level measured 1 in. from the lips and the ear canal. For experimental testing, the vocal signal from the microphone was amplified with a Mackie mixer (model 1202) and shifted in pitch with an Eventide Eclipse Harmonizer, and then amplified with a Crown D75 amplifier and HP 350 dB attenuators. Max/MSP (v. 4.6 by Cycling 74) was used to control the harmonizer. The voice output, feedback, and transistor-transistor logic (TTL) control pulses were low-pass filtered at 5 kHz (finite impulse response filter), digitized at 10 kHz, and recorded on a Macintosh computer using CHART software (AD Instruments, Castle Hill, New South Wales, Australia).

C. Procedure

Subjects were asked to vocalize the vowel sound /a/ for about 5 s. For each vocalization, the pitch feedback was downward shifted a total of five times with a randomized inter-stimulus interval between 700 and 900 ms. The experimental procedure consisted of blocks of 8–16 vocalizations, producing 40–80 pitch-shift stimuli per subject. The stimulus duration was fixed at 200 ms and the magnitude was held constant at –100 cents (100 cents = 1 semitone). Response magnitudes are calibrated in cents, which is a relative logarithmic scale that allows normalization of voice F_0 across subjects. The voice waveform was processed offline in PRAAT (Boersma, 2001) using an autocorrelation method to produce a train of pulses corresponding to the momentary fundamental period of the voice signal recorded via microphone. This pulse train was then converted into an analog wave in IGOR PRO (Wavemetrics, Inc., Lake Oswego, OR). The analog F_0 wave was then converted from Hz to cents with the formula

$$\text{cent} = 12 \times \frac{\log(F_0/196)}{\log 2} \times 100,$$

where F_0 denotes the fundamental frequency in Hz, 196 = arbitrary reference frequency.

For the purpose of obtaining the average F_0 response, voice signals for each trial were time-aligned with the trigger pulse, using 200 ms pre- and 500 ms post-stimulus windows, and averaged. Prior to the averaging, inspections of the individual trials were performed to eliminate those trials with unusually large amplitude, resulting from either signal processing errors or vocal interruption. Valid averaged vocal responses were defined by the averaged cent waveform exceeding a value of 2 standard deviations (SDs) of the pre-stimulus mean (baseline F_0) beginning at least 60 ms after the stimulus and lasting at least 50 ms (Chen *et al.*, 2007). The criterion for the end of a response is that it returns to within 2 SDs of the pre-stimulus mean for at least 30 ms. Latency of the averaged response was defined as the time from the stimulus onset at which the response exceeds 2 SDs of the pre-stimulus mean, and the response magnitude was defined as the difference between the pre-stimulus mean and the greatest value of the F_0 contour following the response onset. Measures of vocal response magnitude and latency were taken from the averaged cent waveforms and those values were analyzed in SPSS (v. 16.0) for significance tests with one-way analyses of variance (ANOVAs) (age).

III. RESULTS

Figure 1 shows representative voice F_0 responses to 100 cents downward stimuli generated by the three groups of subjects. From top to bottom are the response contours for the children group, the younger adult group, and the older group, respectively. As can be seen from Fig. 1, these response contours were upward, as were all other responses in this study, and were in the opposite direction to the stimuli. While the response magnitudes appear to be roughly similar, the response of the child has a longer latency than the young or older adults.

Figure 2 shows boxplots of response magnitudes (top) and latencies (bottom) as a function of age. Table 1 shows the average values and SDs of response magnitudes and latencies across age. One-way factorial ANOVAs were performed on the response magnitude and latency as a function of age. For the response magnitude, significant differences were found among three age groups [$F(2,27)=5.822$, $p=0.014$]. *Post hoc* Bonferroni tests indicated that the older group (33 cents) produced significantly larger response magnitudes than the children group (21 cents) ($p=0.039$) and young adults group (20 cents) ($p=0.026$). Statistical analyses also revealed significant differences in the response latency among three groups [$F(2,27)=0.012$], and *post hoc* Bonferroni tests indicated that the children group (223 ms) produced significantly longer latencies than the young adults group (126 ms) ($p=0.033$) and older adults group (121 ms) ($p=0.024$).

FIG. 1. Representative contours of voice F_0 responses to 100 cents downward stimuli for the children group (top), the younger adults group (middle), and the older adults group (bottom). Horizontal dotted lines indicate ± 2 SDs of the pre-stimulus mean voice F_0 . Vertical dashed lines affixed to averaged F_0 responses indicate onset and offset times of the response. Horizontal dashed line indicates response magnitude.

IV. DISCUSSION

The present study compared responses to pitch-shifted voice feedback in English speakers across three different age groups to test whether age is a contributing factor to the modulation of voice F_0 responses. Significant differences in both the response magnitude and the latency across the age range demonstrate that the online control of voice F_0 by auditory feedback can be affected by the age of the subjects. Furthermore, the greatest response magnitudes for the older group and the longest latencies for the children group suggest age-specific characteristics of the audio-vocal system in voice F_0 control.

Previous research indicates that vocal response latency can be modulated as a function of stimulus magnitude (Larson *et al.*, 2001), voice F_0 level (Liu and Larson, 2007), and vocal task (Chen *et al.*, 2007). The present study provides evidence that the modulation of vocal response latency is age-dependent. The observed prolonged latency of vocal responses in the children group may be regarded as a reflection of typical neurophysiological maturational processes (Rojas *et al.*, 1998; Kotecha *et al.*, 2009). Although direct comparisons in response latency with previous pitch-shift studies are not available, similar results have been found in the studies of magnetoencephalography (MEG) responses in the auditory cortex to an auditory stimulus. It has been demonstrated that auditory evoked response latency (M100) is dependent on age in healthy developing children, with prolonged latencies for children and shorter latencies for those approaching

FIG. 2. Boxplots of averaged response magnitudes and latencies across age. 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).

adulthood (Rojas *et al.*, 1998; Kotecha *et al.*, 2009). Moreover, no significant differences were observed in the response latencies between the younger and the older adults, indicating the absence of an age effect on the response latency in the adults. Therefore, the similar increase in response latency for the pitch-shift reflex and auditory evoked potentials suggests that similar mechanisms may be involved. Namely, if processing of auditory feedback at the brainstem and cortical level, as revealed through MEG studies, is slower in children compared with adults, this would very likely affect the pitch-shift reflex since it is dependent on auditory feedback. These results also suggest that a long latency or slower processing time of the pitch-shift reflex could serve as an effective indicator of auditory function in developing children.

As compared to the children and the younger adults, the older adults produced the largest response magnitudes. No significant differences were observed between the children and the younger adults. These findings suggest that the in-

crease in the response magnitude from the children and the younger adults to the older adults may be attributed to the aging-specific physiological changes. These changes include immunological, neurological, respiratory, gastrointestinal, muscular, dermatological systems, etc. (Ramig *et al.*, 2001). Since the neural and kinesthetic mechanisms underlying auditory feedback control of voice F_0 are still unclear, there is no specific explanation for the greater response magnitudes in the older adults. According to our recent modeling work in auditory feedback control of voice F_0 (Larson *et al.*, 2008), however, it is possible that the age-related changes in neuromuscular, sensory feedback, and motor control might make major contributions for the response magnitudes generated by the older groups. Along these lines, it is possible that, as people age, they become more sensitive to changes in their voice auditory feedback and produce larger responses. If this is the case, it is surprising that such an effect would not also be evident in the transition from childhood to young adult status.

The primary limitation of the present study is that response magnitude and latency were measured from only one type of pitch-shift stimulus (100 cents, downward direction). Other parameters such as direction, magnitude, voice F_0 , and so on should be undertaken in future studies to address the robust effect of age on voice F_0 responses. Another limitation of the current study is the lack of information concerning the gender effect on vocal response in magnitude and latency. Although there is no published report verifying the sex difference in auditory feedback control of voice, sex does impact the acoustic characteristics of voice in children (Nicollas *et al.*, 2008) and older adults (Mueller, 1997; Torre and Barlow, 2009). Thus, it is possible that sex might impact the control of voice F_0 in these two populations and needs to be further explored.

Although this preliminary study indicates an age effect on the auditory feedback control of voice F_0 , there are several additional questions that need to be answered. For example, why did children produce similar response magnitudes to those for the younger adults despite their longer latencies? Similarly, why were only significantly larger response magnitudes observed in the older adults compared to the other two groups? These questions cannot be answered at this time because the functional network of the audio-vocal system underlying these responses is still unknown. Future studies using neural imaging and modeling, combined with perturbed auditory feedback, hold promise for defining this network in greater detail.

V. CONCLUSION

The primary purpose of this study was to determine if age has an effect on the regulation of voice F_0 response to pitch perturbation. The results showed significant age effects on both the response magnitude and the latency. Specifically, the children group produced the longest latencies compared with the younger and older adults, while the largest response magnitudes were generated by the older adults compared with the other two groups. This preliminary study provides evidence that age may play an important role in voice F_0

TABLE I. Averaged response magnitudes and latencies (SD) across age.

Age	Response magnitude (cents)	Response latency (ms)
Children	21 (8)	223 (122)
Younger	20 (9)	126 (54)
Older	33 (11)	121 (36)
Total	25 (11)	157 (90)

control, and the developing brain and the aging physiological changes may impact the processing of auditory feedback.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. 1R01DC006243 and Chinese NSF grant No. 30970965. The authors thank Chun Liang Chan for programming assistance.

- Benjamin, B. J. (1981). "Frequency variability in the aged voice," *J. Gerontol.* **36**, 722–726.
- Benjamin, B. J. (1997). "Speech production of normally aging adults," *Semin Speech Lang* **18**, 135–141.
- Bennett, S. (1983). "A 3-year longitudinal study of school-aged children's fundamental frequencies," *J. Speech Hear. Res.* **26**, 137–141.
- Boersma, P. (2001). "Praat, a system for doing phonetics by computer," *Glott 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**, 3153–3161.
- 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.
- 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.
- 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.
- Glaze, L., Bless, D., Milenkovic, P., and Susser, R. (1988). "Acoustic characteristics of children's voice," *J. Voice* **2**, 312–319.
- Glaze, L., Bless, D., and Susser, R. (1990). "Acoustic analysis of vowel and loudness in children's voice," *J. Voice* **4**, 37–44.
- Jones, J. A., and Keough, D. (2008). "Auditory-motor mapping for pitch control in singers and nonsingers," *Exp. Brain Res.* **190**, 279–287.
- Kawahara, H. (1994). "Interactions between speech production and perception under auditory feedback perturbations on fundamental frequencies," *J. Acoust. Soc. Jpn.* **15**, 201–202.
- 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.
- Kotecha, R., Pardos, M., Wang, Y., Wu, T., Horn, P., Brown, D., Rose, D., deGrauw, T., and Xiang, J. (2009). "Modeling the developmental patterns of auditory evoked magnetic fields in children," *PLoS ONE* **4**, e4811.
- 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., Altman, K. W., Liu, H., and Hain, T. C. (2008). "Interactions between auditory and somatosensory feedback for voice F (0) control," *Exp. Brain Res.* **187**, 613–621.
- Larson, C. R., Burnett, T. A., Bauer, J. J., Kiran, S., and Hain, T. C. (2001). "Comparisons of voice F₀ responses to pitch-shift onset and offset conditions," *J. Acoust. Soc. Am.* **110**, 2845–2848.
- Larson, C. R., Burnett, T. A., Kiran, S., and Hain, T. C. (2000). "Effects of pitch-shift onset velocity on voice F0 responses," *J. Acoust. Soc. Am.* **107**, 559–564.
- Liss, J. M., Weismer, G., and Rosenbek, J. C. (1990). "Selected acoustic characteristics of speech production in very old males," *J. Gerontol.* **45**, 35–45.
- 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., Wang, E. Q., Metman, L. V., and Larson, C. R. (2008). "Vocal responses to loudness- and pitch-shift perturbations in individuals with Parkinson's disease," in *Proceedings of the Motor Speech Conference*, Monterey, CA.
- Liu, H., Xu, Y., and Larson, C. R. (2009). "Attenuation of vocal responses to pitch perturbations during Mandarin speech," *J. Acoust. Soc. Am.* **125**, 2299–2306.
- Lombard, E. (1911). "Le signe de l'élevation de la voix (The sign of a rising voice)," *Ann Mal Oreille Larynx* **37**, 101–119.
- Mueller, P. B. (1997). "The aging voice," *Semin Speech Lang* **18**, 159–169.
- 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.
- Nicollas, R., Garrel, R., Ouaknine, M., Giovanni, A., Nazarian, B., and Triglia, J. M. (2008). "Normal voice in children between 6 and 12 years of age: Database and nonlinear analysis," *J. Voice* **22**, 671–675.
- Ramig, L. A., and Ringel, R. L. (1983). "Effects of physiological aging on selected acoustic characteristics of voice," *J. Speech Hear. Res.* **26**, 22–30.
- Ramig, L. O., Gray, S., Baker, K., Corbin-Lewis, K., Buder, E., Luschei, E., Coon, H., and Smith, M. (2001). "The aging voice: A review, treatment data and familial and genetic perspectives," *Folia Phoniatr Logop* **53**, 252–265.
- Rojas, D. C., Walker, J. R., Sheeder, J. L., Teale, P. D., and Reite, M. L. (1998). "Developmental changes in refractoriness of the neuromagnetic M100 in children," *NeuroReport* **9**, 1543–1547.
- Russo, N., Larson, C., and Kraus, N. (2008). "Audio-vocal system regulation in children with autism spectrum disorders," *Exp. Brain Res.* **188**, 111–124.
- Sataloff, R. T., Rosen, D. C., Hawkshaw, M., and Spiegel, J. R. (1997). "The aging adult voice," *J. Voice* **11**, 156–160.
- Schneider, B., Zumtobel, M., Prettenhofer, W., Aichstill, B., and Jocher, W. (2009). "Normative voice range profiles in vocally trained and untrained children aged between 7 and 10 years," *J. Voice* (in press).
- Shipp, T., and Hollien, H. (1969). "Perception of the aging male voice," *J. Speech Hear. Res.* **12**, 703–710.
- Torre, P., and Barlow, J. A. (2009). "Age-related changes in acoustic characteristics of adult speech," *J. Commun. Disord.* **42**, 324–333.
- Wilcox, K. A., and Horii, Y. (1980). "Age and changes in vocal jitter," *J. Gerontol.* **35**, 194–198.
- 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.