ICIC Express Letters
Part B: Applications
Volume **9**, Number **4**, April **2018**

ICIC International ©2018 ISSN 2185-2766
pp. 331–338

INFLUENCE OF THE LOMBARD EFFECT, FLETCHER EFFECT AND BAND-EMPHASIZED AUDITORY FEEDBACK ON SINGING VOICE

Satoshi Iijima, Shunsuke Ishimitsu and Masashi Nakayama

Graduate School of Information Science
Hiroshima City University
3-4-1, Ozuka-Higashi, Asa-Minami-Ku, Hiroshima 731-3194, Japan
ishimitu@hiroshima-cu.ac.jp

Received October 2017; accepted January 2018

Abstract. *In this study, we conducted three experiments to investigate the Lombard effect, the Fletcher effect and the effect of band-emphasized auditory feedback on singing voice. In experiment 1, participants sang while listening either to no noise or 60, 70, 80 or 90 dB of noise. In experiment 2, participants sang while hearing their own voices amplified to 75, 85 or 95 dB. In experiment 3, participants sang while hearing their own voices with the low frequencies removed by a 1,000, 2,000 or 3,000 Hz high-pass filter. From experiment 1, we discovered that the sound pressure level of their voices increased with noise. The results of experiment 2 indicated that the sound pressure level and the first two formant frequencies decreased significantly as the feedback increased. Finally, experiment 3 showed that the pitch and sound pressure level of their voices decreased as the cut-off frequency increased.*

Keywords: Auditory feedback, Singing voice, Lombard effect, Fletcher effect, Band enhancement

1. **Introduction.** Many studies have discovered that the speech process involves both its production and perception. In speech production, people perceive their own voices through their auditory organs and then modulate them through feedback to their vocal organs. This process is called auditory feedback. However, while there have been many studies on auditory feedback for speech, there has not been much research on singing voice.

Lombard demonstrated that the sound pressure level (SPL) of a person's speaking voice increases in a noisy environment [1]. This phenomenon is called the Lombard effect. After this finding, several studies investigated the properties of this effect. It has been shown that the fundamental frequency (F_0), 1st and 2nd formant frequencies (F_1 and F_2), vocal power spectrum and speech duration all increase in the presence of noise [2,3]. On the other hand, when people hear a loud feedback of their own voice, their speech level reduces. This is called the Fletcher effect [4]. It has been reported that the vocal SPL decreases by approximately 0.3-0.6 dB for every 1.0 dB of amplified voice feedback [5,6]. These studies were concerned with conversation. Bottalico et al. measured SPL and singing power ratio produced by nonprofessional and professional singers (10 males and 10 females) during a performance of an excerpt from the Star-Spangled Banner with three levels of the accompaniment (70, 80, and 90 dBA) and with three different levels of external auditory feedback: normal, with reflective panels (increased external auditory feedback), and with headphones (cancelled external auditory feedback) conditions [7]. Bottalico et al. indicated that the SPL increased under the loud accompaniment and higher levels of external auditory feedback were associated with a reduction in SPL.

It is believed that a singer's formant typically appears in their operatic voice [8], at the peak of the spectral envelope observed at around 2.3-3.8 kHz for male singers. Vocal tract tuning is a strategy used for generating a singer's formant [9]. Such tuning has also been found in some other traditional singing voices [10]. It is, therefore, believed that the tuning of formant frequencies is important for singing voice.

In the present study, we performed three experiments to investigate the effects of auditory feedback on singing voice. In the first experiment, we investigated the Lombard effect by observing how the response to noise increased across the range of 60-90 dB in 10 dB steps. In the second experiment, we examined the Fletcher effect by comparing the effects caused by voice feedback at the same level and 10 dB more and 10 dB less than the level of the vocal utterance. In the third experiment, we used a high-pass filter to observe people's responses to feedback that emphasized the high frequencies of their voices. Here, we consider a voice to be a singing voice if specific vowels are uttered at specific pitches. It is thought that singing is a different type of talking. The differences between talking and singing are that singing voices change in pitch and produce utterances with longer durations than talking voices.

In this paper, we reviewed some of the early studies as Introduction. Then, we showed our experimental methods in the next chapter. After that, we showed some of the data taken from these experiments and discussed about the results. Finally, we summarized the present study and indicated our future plans.

2. Methods.

Experiment 1 investigated responses to stepwise changes in noise levels by examining the effects of listening to noise via headphones on people's singing. Experiment 2 assessed responses to stepwise changes in voice feedback levels to clarify how changing the amplitude of a singer's own voice affects him/herself. Experiment 3 investigated responses to stepwise changes in the cut-off frequencies of band-emphasized voice feedback to observe the role of higher frequencies of a singer's own voice in auditory feedback.

2.1. Participants.

Six healthy young males (aged 21-23) participated in experiments 1, 2, and 3. Five of the participants also took part in all the experiments. They were first required to pass a test for their sense of pitch. Two successive tones, either the same, minor second, perfect fifth or octave apart, were generated using a keyboard and the participants were asked whether the second tone was higher, lower or the same as the first tone. None of the participants reported a history of neurological, speech or hearing disorders, and none had received any professional vocal training.

2.2. Apparatus.

The participants wore headphones (SONY/MDR-Z7) in an anechoic chamber during the experiment. They were asked to vocalize at 85 dB while looking at a sound level meter (RION/NL-31). A-weighting was used as the measurement method. The vocal signal was fed from a microphone (Brüel & Kjær/Type 4189) to a microphone pre-amplifier (Brüel & Kjær/Type 2671) and then to a microphone amplifier unit (ONOSOKKI/SR-2200). The noise and voice feedback was passed through a mixer (ZOOM/R24) in experiment 1. We used pink noise, which was amplified gradually in 10 dB steps from 60 to 90 dB and was presented to the participants via a PC (MacBook Air). The voice feedback was amplified to 85 dB to maintain a level similar to the original utterance and avoid the Fletcher effect. The SPL (A-weighted) was measured at the headphone ear pad with a sound level meter (Brüel & Kjær/Type 2250). In experiment 2, the voice feedback was amplified gradually from 75 to 95 dB in 10 dB steps. In experiment 3, we used a high-pass filter (NF Corporation/MS-525) with cut-off frequencies of 1,000, 2,000 and 3,000 Hz. The high-pass filter used was an eighth-order Butterworth filter. Frequencies lower than the cut-off frequencies were attenuated by 48 dB per octave by the high-pass filter. The high-pass-filtered voice feedback was then amplified to 85 dB, causing sounds above the cut-off frequencies to be emphasized. No masking noise was used to mask the participants' air- and bone-conducted sound in experiments 2 and 3.

Recording was performed at a sampling frequency of 65,536 Hz with 24-bit quantization by a digital recorder (Brüel & Kjær/LAN-XI3050-060).

2.3. **Procedure.** The participants were instructed to sing a steady vowel /a/ at three pitches, C3 (130.81 Hz), G3 (196 Hz) and C4 (261.63 Hz), for approximately 5 s in experiments 1, 2, and 3. Before the experiment, several practice trials were conducted to ensure that the participants could match the notes within 100 cent. Sine waves, at pitches double those of C3, G3 and C4, were presented to the participants using the PC before each trial. After the experimenter gave the signal, the participants began to sing. We doubled the frequencies of the sine waves because the actual sounds were so low-frequency that it was difficult to confirm their pitches. While the participants were singing, the sine waves were not presented. In experiment 1, the noise conditions were no noise and 60, 70, 80 and 90 dB pink noise. Each experimental block included 15 vocalizations. Each note was repeated for each of the four noise levels and the no-noise condition. The trials were randomized and performed five times each for a total of 75 trials per participant. In experiment 2, the voice feedback was presented at 75, 85 and 95 dB via headphones. Each experimental block included nine vocalizations. Each note was repeated with voice feedback at each of the three feedback levels. The trials were randomized and performed five times each, for a total of 45 trials per participant. In experiment 3, the voice feedback was presented through the high-pass filter via headphones. Each note was repeated with high-pass-filtered feedback at each of the three cut-off frequencies and with no filtering. The trials were randomized and performed five times each, for a total of 60 trials per participant.

2.4. **Data analysis.** We obtained 450 valid trials (6 participants $\times$ 5 noise conditions $\times$ 3 pitches $\times$ 5 sets) in experiment 1, 270 valid trials (6 participants $\times$ 3 feedback conditions $\times$ 3 pitches $\times$ 5 sets) in experiment 2 and 360 valid trials (6 participants $\times$ 4 band-emphasized conditions $\times$ 3 pitches $\times$ 5 sets) in experiment 3. Praat [11] was used to analyze the voice waveforms for each participant separately in terms of F_0, SPL and F_1 and F_2 for each experimental condition in each experiment. The formants were analyzed using the Burg method with a maximum of five formants. In the analysis, we used the vocal signals that were recorded from 0.5 s to 3.5 s after the beginning of the utterance and averaged the data for each condition. The data were subjected to a two-way analysis of variance (ANOVA) (IBM SPSS, v. 17.0) for three pitches and five noise levels in experiment 1, three pitches and three voice feedback levels in experiment 2 and three pitches and four cut-off frequencies in experiment 3. A Tukey *post hoc* test was carried out where there were significant differences. We also executed a Tukey-Kramer test (JMP, v. 13) for each pitch with respect to noise in experiment 1, feedback level in experiment 2 and cut-off frequency in experiment 3.

3. Results.

3.1. Experiment 1.

3.1.1. *Fundamental frequency.* Here, the fundamental frequencies are expressed by converting to cent from Hz (100 cent = 1 semitone). A two-way ANOVA analysis showed that the main effect of pitch was significant: $F(2, 435) = 35.55$, $p < .0001$. A Tukey *post hoc* test revealed that C3 was 37.36 and 32.03 cent nearer the target pitch than G3 and C4 respectively ($p < .0001$). The main effect of noise and the interaction effect were not significant: $F(4, 435) = 0.33$, $p = .86$ and $F(8, 435) = 0.22$, $p = .99$.

3.1.2. *Sound pressure level.* A two-way ANOVA analysis showed that the main effect of pitch was significant ($F(2, 435) = 14.89$, $p < .0001$), as was that of noise ($F(4, 435) = 4.04$, $p = .003$). C3 was 1.15 and 1.82 dB lower than G3 and C4: $p = .002$ and $p < .0001$, as well as no noise was 1.48 dB and 60 dB noise was 1.38 dB lower than 90 dB noise ($p = .007$ and $p = .015$). The interaction effect was not significant: $F(8, 435) = 0.13$, $p = .998$.

3.1.3. *Formant frequency.* A two-way ANOVA analysis showed that the main effect of pitch was significant for F_1: $F(2, 435) = 60.78$, $p < .0001$. C3 was 22.86 and 64.35 Hz lower than G3 and C4 respectively ($p < .0001$) and G3 was 41.49 Hz lower than C4 ($p < .0001$). The main effect of noise and the interaction effect were not significant: $F(4, 435) = 1.42$, $p = .23$ and $F(8, 435) = 0.27$, $p = .97$. For F_2, the main effect of pitch was significant: $F(2, 435) = 182.66$, $p < .0001$. C3 was 104.04 and 164.59 Hz lower than G3 and C4 respectively ($p < .0001$) and G3 was 60.55 Hz lower than C4 ($p < .0001$). The main effect of noise and the interaction effect were not significant: $F(4, 435) = 0.17$, $p = .95$ and $F(8, 435) = 0.12$, $p = .998$.

3.2. **Experiment 2.**

3.2.1. *Fundamental frequency.* A two-way ANOVA analysis showed that the main effect of pitch was significant: $F(2, 261) = 7.81$, $p = .001$. A Tukey *post hoc* test revealed that C3 and G3 were 15.50 and 13.90 cent nearer the target pitch than C4: $p = .001$ and $p = .004$. The main effect of feedback level and interaction effect were not significant: $F(2, 261) = 0.81$, $p = .44$ and $F(4, 261) = 0.63$, $p = .64$.

3.2.2. *Sound pressure level.* A two-way ANOVA analysis showed that the main effects of pitch and feedback level were significant: $F(2, 261) = 35.63$, $p < .0001$ and $F(2, 261) = 49.11$, $p < .0001$. C3 was 1.49 and 1.70 dB lower than G3 and C4 ($p < .0001$), as well as between feedback levels of 75 were 0.73 and 2.15 dB higher than levels of 85 ($p = .003$) and 95 ($p < .0001$) and level 85 was 1.42 dB higher than 95 ($p < .0001$). The interaction effect was not significant: $F(4, 261) = 0.71$, $p = .59$. The SPL decreased as the feedback level was increased for each note (Figure 1). We found that level 75 was 1.69 dB higher than 95 ($p = .002$) and 85 was 1.28 dB higher than 95 ($p = .028$) for C3; level 75 was 1.03 and 2.31 dB higher than 85 ($p = .004$) and 95 ($p < .0001$), 85 was 1.29 dB higher than 95 ($p = .0003$) for G3 and level 75 was 2.32 dB higher than 95 ($p < .0001$) and 85 was 1.66 dB higher than 95 ($p < .0001$) for C4.

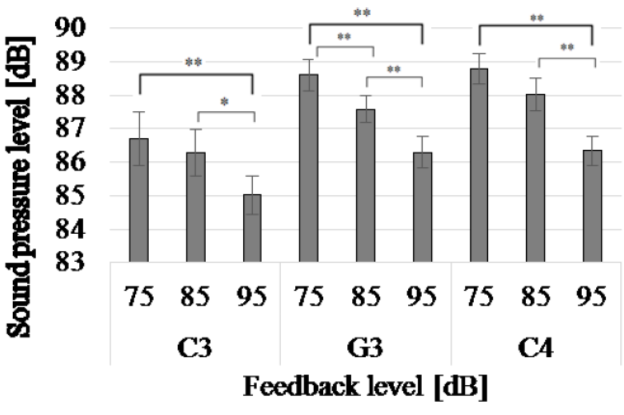

FIGURE 1. Mean SPLs for the three feedback conditions for each note in experiment 2. The error bars represent 95% confidence intervals. The sets of three bars from left to right represent the feedback conditions for each target pitch. The difference significances were calculated using a Tukey-Kramer test (*$p < .05$, **$p < .01$).

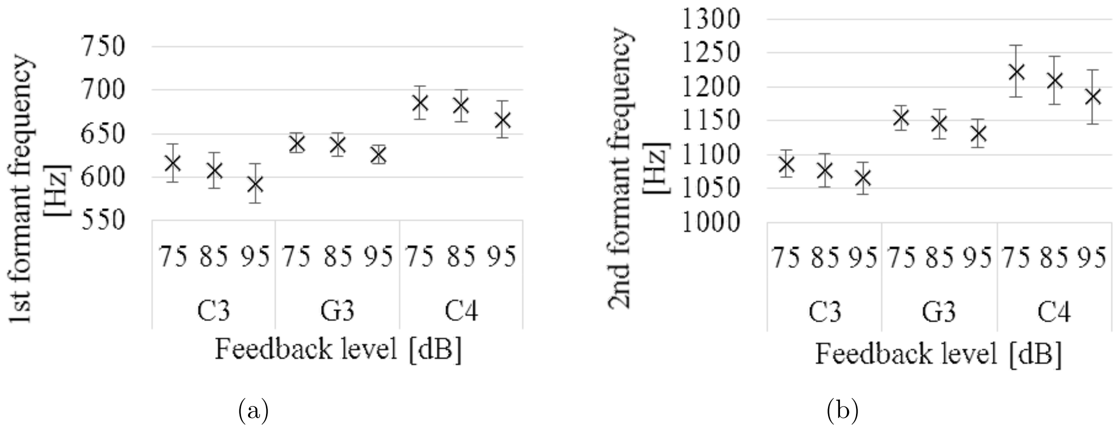

Figure 2. Mean values of F_1 (a) and F_2 (b) for the three feedback conditions for each note in experiment 2. The error bars represent 95% confidence intervals. The sets of three crosses from left to right represent the feedback conditions for each target pitch.

3.2.3. *Formant frequency.* A two-way ANOVA analysis showed that the main effects of pitch and feedback level were significant for F_1: $F(2, 261) = 48.21$, $p < .0001$ and $F(2, 261) = 3.43$, $p = .03$. C3 was 28.87 and 72.47 Hz lower than G3 and C4 ($p < .0001$) and G3 was 43.60 Hz lower than C4 ($p < .0001$), as well as feedback level of 75 was 18.75 Hz lower than 95 ($p = .03$). The interaction effect was not significant: $F(4, 261) = 0.90$, $p = .99$. The main effects of pitch and feedback level were significant for F_2: $F(2, 261) = 65.30$, $p < .0001$ and $F(2, 261) = 2.95$, $p = .05$. C3 was 67.48 and 129.74 Hz lower than G3 and C4 ($p < .0001$), G3 was 62.27 Hz lower than C4 ($p < .0001$), as well as feedback level of 75 was 27.31 Hz lower than 95 ($p = .04$). The interaction effect was not significant: $F(4, 261) = 0.11$, $p = .98$.

3.3. **Experiment 3.**

3.3.1. *Fundamental frequency.* A two-way ANOVA analysis showed that the main effects of pitch and band enhancement were significant: $F(2, 348) = 16.55$, $p < .0001$ and $F(3, 348) = 5.00$, $p = .002$. A Tukey *post hoc* test revealed that C3 was 19.49 and 27.85 cent nearer the target pitch than G3 and C4 ($p < .0001$), as well as no filtering were 17.40 and 19.96 cent nearer the target pitch than 2,000 and 3,000 Hz high-pass filtering ($p = .01$ and $p = .003$) respectively. The interaction effect was not significant: $F(6, 348) = 1.59$, $p = .15$. The G3 pitch was 31.10 cent lower for a cut-off frequency of 1,000 Hz than for the unfiltered condition ($p = .03$).

3.3.2. *Sound pressure level.* A two-way ANOVA analysis showed that the main effects of pitch and band enhancement were significant: $F(2, 348) = 8.69$, $p < .0001$ and $F(3, 348) = 23.96$, $p < .0001$. C3 was 0.87 and 0.79 dB lower than G3 ($p = .0001$) and C4 ($p = .002$), as well as no filtering were 1.63 and 1.81 dB higher than 2,000 and 3,000 Hz high-pass filtering ($p < .0001$) and 1,000 Hz were 1.36 and 1.54 dB lower than 2,000 and 3,000 Hz high-pass filtering ($p < .0001$). The interaction effect was not significant: $F(6, 348) = 0.65$, $p = .69$. The SPL decreased as the cut-off frequency increased for each note (Figure 3). For C3, no filtering was 1.35 dB higher than 3,000 Hz ($p = .03$) and 1,000 Hz was 1.41 dB higher than 3,000 Hz high-pass filtering ($p = .02$). For G3, no filtering was 1.86 and 1.81 dB higher than 2,000 and 3,000 Hz ($p = .0006$ and $p = .0009$) and 1,000 Hz was 1.24 dB higher than 2,000 Hz high-pass filtering ($p = .04$). For C4, no filtering was 1.95 and 2.29 dB higher than 2,000 and 3,000 Hz ($p = .0002$ and $p < .0001$)

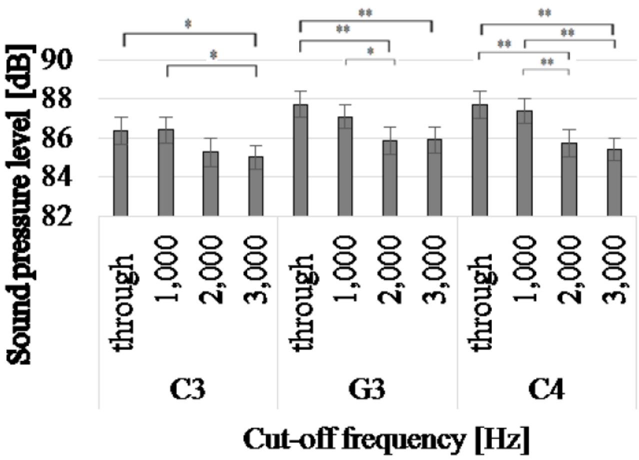

Figure 3. Mean SPLs for the four high-pass-filtered conditions for each note in experiment 3. The error bars represent 95% confidence intervals. The sets of four lines from left to right represent the high-pass filtering conditions for each target pitch. The difference significances were calculated using a Tukey-Kramer test (*$p < .05$, **$p < .01$).

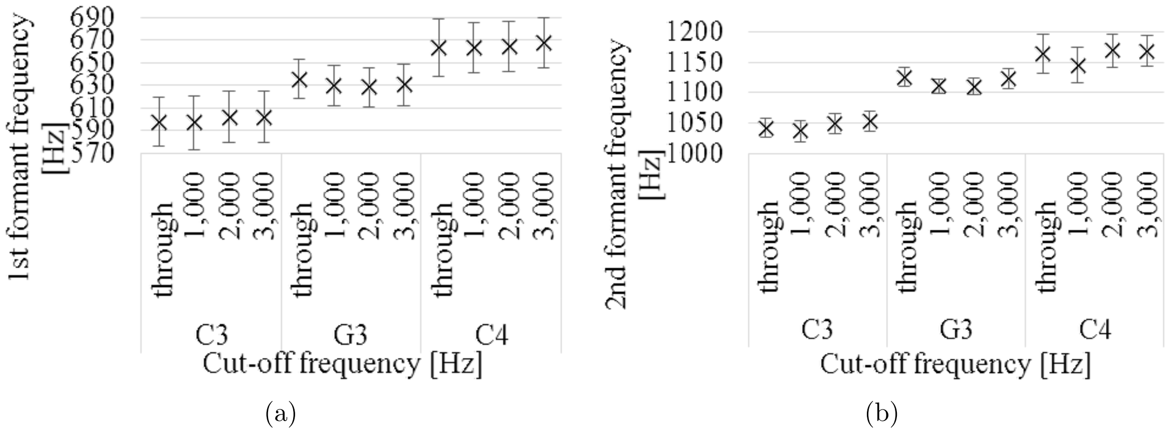

Figure 4. Mean frequencies of F_1 (a) and F_2 (b) for the four high-pass-filtered conditions for each note in experiment 3. The error bars represent 95% confidence intervals. The sets of four crosses from left to right represent the noise conditions for each target pitch.

and 1,000 Hz were 1.67 and 2.01 dB higher than 2,000 and 3,000 Hz high-pass filtering ($p = .002$ and $p = .0001$).

3.3.3. *Formant frequency.* A two-way ANOVA analysis showed that the main effect of pitch was significant for F_1: $F(2, 348) = 37.78$, $p < .0001$. C3 was 31.31 and 64.95 Hz lower than G3 and C4 ($p < .0001$). G3 was 33.64 Hz lower than C4 ($p < .0001$). The main effect of band enhancement and the interaction effect were not significant: $F(3, 348) = 0.05$, $p = .99$ and $F(6, 348) = 0.07$, $p = .999$.

The main effect of pitch was significant for F_2: $F(2, 348) = 129.74$, $p < .0001$. C3 was 71.64 and 116.13 Hz lower than G3 and C4 ($p < .0001$) and G3 was 44.49 lower than C4 ($p < .0001$). The main effect of band enhancement and the interaction effect were not significant: $F(3, 348) = 1.56$, $p = .20$ and $F(6, 348) = 0.38$, $p = .90$. There were no significant differences in the F_1 and F_2 values under the different experimental conditions (Figure 4): we obtained almost the same results for each note under all different conditions.

4. **Discussion.**

4.1. **Fundamental frequency.** We found more stable pitches in singing than in speech in experiments 1 and 2. In speech production, increases in F_0 under noise have been widely reported [1-3]. However, we did not find the same increases in F_0 in experiment 1. In singing, people are required to match their voices to the target pitches, which is believed to be the reason why it is difficult to affect F_0 in singing even if people are exposed to noise.

Similarly, the results for F_0 in experiment 2 were almost the same as those in experiment 1. On the other hand, we observed decreases in F_0 when participants heard band-emphasized auditory feedback with a cut-off frequency of over 2,000 Hz in experiment 3. Vowel information is not included in vocal sound above 2,000 Hz. The results of experiment 3 indicate that there are differences in the effect depending on whether vowel information is included.

These results suggest that loud noise and quiet or loud voice feedback do not affect F_0 in singing, but when people hear their own voices with frequencies above 2,000 Hz emphasized, their F_0 may decrease. However, further research may be required with regard to F_0 decreasing in this case.

4.2. **Sound pressure level.** It is also well known that vocal SPLs increase in noisy environments. In experiment 1, we observed a tendency for the SPL to increase with noise. In previous reports [1-3,7], the speech SPLs were freely chosen, so it was thought that these increases in SPL with noise were natural. However, in the present study, where the participants uttered a vowel sound while watching the sound level meter, their SPLs still increased as the noise increased, so we consider that the loudness of the noise has a highly significant influence. Similar results were obtained in experiments 2 and 3. The SPLs decreased with greater feedback as the cut-off frequencies became higher. These results indicate that the feedback levels of the participants' own voices had a strong influence on the SPLs of their utterances.

4.3. **Formant frequency.** Although it has been thought that the values of F_1 and F_2 increase under noise [2], the results of experiment 1 do not support this statement. In experiment 2, the F_1 and F_2 values decreased as the feedback level increased. In experiment 3, the F_1 value was constant regardless of the cut-off frequency. Although there were differences between the results of experiments 2 and 3 for F_1, the SPL results show a common trend, as can be seen in Figures 1 and 3. The differences in these response sounds were dictated by whether the low frequencies were cut from the vocal sound. In experiment 3, when the high-pass filter was used, it removed the participants' F_1. As they did not hear their F_1 loudly, they maintained their F_1 value. Therefore, it is possible that a lack of lower frequencies affects the production and perception of singing voice. Our hypothesis is that, when people receive loud feedback of their own voice that includes F_1, they unconsciously feel that their F_1 is sufficiently loud and so decrease their F_1 automatically. This would explain why the participants maintained their F_1 in experiment 3.

5. **Conclusions.** The present study has investigated the Lombard effect, the Fletcher effect and the effect of band-emphasized auditory feedback on singing voice. The F_0 value was not affected in the Lombard and Fletcher environments, but decreased when the participants received voice feedback that was emphasized above 2,000 Hz. As the feedback noise increased, a trend for the SPL of the participant's voice to increase was observed. When the participants heard their own voice loudly, their SPL decreased significantly. We also observed that, when the participants heard their voice with the higher frequencies emphasized, their SPL decreased significantly. Finally, the results indicate

that if a sufficiently high F_1 value is perceived by the auditory organs, the vocal response will be decreased. However, further research is needed about the role of higher frequency in auditory feedback. In future work, we will investigate the band-emphasized auditory feedback with other situations such as changing the emphasized band and feedback level.

Acknowledgment. This work was supported by the Kawai Foundation for Sound Technology & Music.

REFERENCES

[1] E. Lombard, Le signe de l'elevation de la voix, *Annuals Maladies Oreille Larynx*, Nez, Pharynx, vol.37, pp.101-109, 1911.
[2] Y. Uemura, M. Morise and T. Nishiura, Improvement of speech recognition performance based on the conversion of Lombard features, *IEICE Technical Report*, SP2010-1, pp.1-6, 2010.
[3] R. Patel and K. W. Schell, The influence of linguistic content on the Lombard effect, *Journal of Speech, Language, and Hearing Research*, vol.51, pp.209-220, 2008.
[4] H. Fletcher, G. M. Raff and F. Parmley, Study of the effects of different sidetones in the telephone set, *Report (19412)*, Western Electrical Company, 1918.
[5] H. Lane and B. Tranel, The Lombard sign and the role of hearing in speech, *Journal of Speech, Language, and Hearing Research*, vol.14, no.4, pp.677-709, 1971.
[6] G. M. Siegel and H. L. Pick Jr, Auditory feedback in the regulation of voice, *The Journal of the Acoustical Society of America*, vol.56, no.5, pp.1618-1624, 1974.
[7] P. Bottalico, S. Graetzer and E. J. Hunter, Effect of training and level of external auditory feedback on the singing voice: Volume and quality, *Journal of Voice*, vol.30, no.4, pp.434-442, 2016.
[8] J. Sundberg, *The Science of the Singing Voice*, Northern Illinois University Press, Dekalb, IL, 1987.
[9] J. Sundberg, F. M. Lã and B. P. Gill, Formant tuning strategies in professional male opera singers, *Journal of Voice*, vol.27, no.3, pp.278-288, 2013.
[10] N. Henrich, M. Kiek, J. Smith and J. Wolfe, Resonance strategies used in Bulgarian women's singing style: A pilot study, *Logopedics Phoniatrics Vocology*, vol.32, no.4, pp.171-177, 2007.
[11] P. Boersma and D. Weenink, *Praat*, http://www.fon.hum.uva.nl/praat/.