24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

EFFECTS OF AUDITORY FEEDBACK ON SINGING

Satoshi Iijima, Shunsuke Ishimitsu, Masashi Nakayama

Graduate School of Information Science, Hiroshima City University, 3-4-1 Ozuka-Higashi, Asa-Minami-Ku, Hiroshima, 7313194 Japan,
e-mail: ishimitu@hiroshima-cu.ac.jp

When people speak in noisy environments, the fundamental frequency, sound pressure level, formant frequency, and duration of their voices increase. On the other hand, if people speak while receiving their own voice loudly, the intensity of their voices decreases. These phenomena, which are known as the Lombard effect and Fletcher effect, respectively, have been explored to clarify the role of auditory feedback on speech. In this study, we conducted three experiments to investigate these two effects and the effect of high-pass filtered auditory feedback on singing. In experiment 1, participants uttered the vowel /a/ at 85 dB for approximately 5 s while listening to no noise, followed by 60, 70, 80, and 90 dB of pink noise at three pitch heights: C3, G3, and C4. In experiment 2, the participants received their own voices, amplified to 75, 85, and 95 dB, while they sang under the same conditions as the first experiment. In the third experiment, the participants uttered while receiving their own voice with the low frequencies of their voices cut off at 1,000, 2,000, and 3,000 Hz with a high-pass filter under the same conditions as the former experiments. As a result of experiment 1, we discovered the possibility that the sound pressure level of their voices increased as the noise increased. The result of experiment 2 indicated that the sound pressure level decreased significantly and showed the potential for a tendency to decrease the first two formant frequencies. Additionally, the result of experiment 3 showed that the pitch and the sound pressure level of their voices decreased as the cut-off frequency increased. However, while the formant frequency decreased in the results of experiment 2, it was constant in experiment 3. This result indicates the possibility that, if a sufficiently high level of formant frequency was perceived, it was decreased sensitively.

Keywords: Auditory feedback, Singing, Lombard effect, Fletcher effect, Emphasized high-frequency

1. Introduction

There are many studies that reveal that the speech process includes its production and perception. In speech production, people receive their own voices through their auditory organs and then modulate them through feedback to their vocal organs. This feedback process is called auditory feedback. However, while there are many studies on auditory feedback of speech, there is not much research on singing.

Lombard demonstrated that the sound pressure level of the speaking voice of a person increases in a noisy environment [1]. This phenomenon is called the Lombard effect. After this finding, several studies have proven the features of this effect. It was shown that the vocal sound pressure level increased by approximately 0.4–0.5 dB for every 1.0 dB increase in the noise level up to 50–90 dB [2, 3]. Moreover, it was found that the fundamental frequency (F_0), 1st and 2nd formant frequencies (F_1 and F_2), power spectrum of the voice, and speech duration increased under noise [4, 5]. On the other hand, when people receive a loud feedback of their own voice, their speech level becomes low. This is called the Fletcher effect [6]. It has been reported that the vocal sound pressure level decreased by approximately 0.3–0.6 dB for every 1.0 dB of amplified feedback voice [7, 8]. These studies were concerned with conversation. There has been little research for singing on these effects.

It is said that the singer's formant appears typically in the operatic voice [9]. It is the peak of the spectral envelope observed at around 2.3‒3.8 kHz for male singers. It has been found that vocal tract tuning is a strategy used in generating the singer's formant [10]. For example, the F_1 of vowel /a/ appears at around 600 Hz. Sopranos often sing higher than 600 Hz in F_0. It is frequently found that there is a situation where the value of F_1 is lower than that of F_0. Hence, to avoid the situation, they increase F_1 to close in on the value of F_0 [11]. This kind of tuning has been found in some other traditional singing voices [12]. Therefore, it is considered that the tuning of formant frequencies is important for a singing voice.

In the present study, we performed three experiments to investigate the effects of auditory feedback on a singing voice. In the first experiment, we investigated the Lombard effect to observe responses caused by noise increased across the range of 60‒90 dB at 10 dB intervals. In the second experiment, we examined the Fletcher effect to compare effects caused by voice feedback levels the same, 10 dB greater or 10 dB smaller than the utterance voice level. In the third experiment, we used a high-pass filter to observe the response when people receive a feedback that emphasized the high frequencies of their voices. In the present study, we have considered the voice uttered in a specific vowel at specific pitch heights as singing. It is thought that singing is a different type of talking. The difference between talking and singing are that a singing voice has a change of pitch and duration longer than a talking voice.

2. Methods

Experiment 1 investigated responses to stepwise shifted noise levels to reveal the effects of people singing under noise. Experiment 2 assessed responses to stepwise shifted voice feedback levels to clarify how the amplitude of a singer's own voice affects him/herself. Experiment 3 investigated responses to stepwise shifted cut-off frequencies of high-pass filters of feedback voices to observe the role of the higher frequencies of a singer's own voice on auditory feedback.

2.1 Experiment 1: Effects of stepwise shifted noise on singing.

2.1.1 Participants

Six healthy young males (ages 21–23) participated in experiment 1. They were first required to pass a test for their sense of pitch. Two successive tones containing unison, minor second, perfect fifth, and octave were generated using a keyboard and the participants were asked to answer whether the second tone was higher, lower, or the same compared to the first tone. None of the participants reported a history of neurological, speech, or hearing disorders. None of them had received any professional vocal training.

2.1.2 Apparatus

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 a measurement method. The vocal signal from a microphone (Brüel & Kjær / Type 4189) and a microphone preamplifier (Brüel & Kjær / Type 2671) was amplified with a microphone amplifier unit (ONOSOKKI / SR-2200). The noise and the feedback voice was passed through a mixer (ZOOM / R24). We used pink noise as the noise. The noise was amplified gradually at every 10 dB intervals from 60 to 90 dB and was presented via a PC (MacBook Air). The feedback voice was amplified to 85 dB to obtain a level similar to the utterance in order to avoid the Fletcher effect. The sound pressure level (A-weighted) at the ear pad of the headphones was measured with a sound level meter (Brüel & Kjær / Type 2250). Recording was performed with a sampling frequency of 65,536 Hz and a quantization bit number of 24 bit by the digital recorder (Brüel & Kjær / LAN-XI3050-060).

2.1.3 Procedures

In the first experiment, participants were instructed to sing a steady vowel /a/ at three pitch heights, C3 (130.81 Hz), G3 (196 Hz), and C4 (261.63 Hz), for approximately 5 s. Before the experiments, several practice trials were conducted to ensure that the participants could match the notes within 100 cent. Sine waves, the pitch heights of which were twice those of C3, G3, and C4, were presented using the PC before each trial. After the experimenter gave the signal, participants began the utterance. We doubled the frequency of the sine waves because the actual sounds were so low that it was difficult to confirm their pitch heights. During the utterance, the sine waves were not presented. The noise conditions were no noise and 60, 70, 80, and 90 dB pink noise. In experiment 1, each experimental block included 15 vocalizations. For each note, the utterance conditions were comprised of the four noise levels and the no-noise condition. These trials were randomized and performed five times each for a total of 75 trials per participant.

2.2 Experiment 2: Effects of stepwise shifted amplitude of singer's own voice.

2.2.1 Participants

Six healthy young males (ages 21–23) took part in experiment 2. Five of the participants also took part in experiment 1. Participants were deemed suitable using the same requirements as experiment 1.

2.2.2 Apparatus

The apparatus and settings were similar to experiment 1. In experiment 2, the feedback voice was amplified gradually from 75 to 95 dB at 10 dB intervals. A masking noise to mask the participants' air and bone conducted sound was not used.

2.2.3 Procedures

In the same manner as the first experiment, participants uttered a steady vowel /a/ at three pitch heights, C3, G3, and C4, for approximately 5 s. The feedback voice was presented at 75, 85, and 95 dB via headphones. In experiment 2, each experimental block included 9 vocalizations. For each note, the voice feedback was comprised of the three feedback levels. These trials were randomized and performed five times each, for a total of 45 trials per participant.

2.3 Experiment 3: Effects of high-pass filtered auditory feedback on singing.

2.3.1 Participants

Six healthy young males (ages 21–23) joined experiment 3. Five of these subjects participated in both experiment 1 and 2. The participants were deemed suitable under the same requirements as the former experiments.

2.3.2 Apparatus

The apparatus and settings were similar to those used in experiments 1 and 2. In experiment 3, we used a high-pass analogue filter (NF Corporation MS-525) with a cut-off frequency of 1,000, 2,000, and 3,000 Hz. The characteristic of the high-pass filter 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 feedback voices were amplified to 85 dB. Amplifying the high-pass filtered voice to 85 dB causes sounds over cut-off frequencies to be emphasized. Any masking noise to mask the participants' air and bone conducted sound was not used.

2.3.3 Procedures

Participants uttered a steady vowel /a/ at three pitch heights, C3, G3, and C4, for approximately 5 s as per the former experiments. The feedback voice was presented through a high-pass filter via headphones. In experiment 3, each experimental block included 12 vocalizations. For each note, the voice feedback was comprised of the three feedback levels and no high-pass filtered condition. These trials were randomized and performed five times each for a total of 60 trials per participant.

2.3.4 Data analysis

We obtained 450 valid trials in experiment 1, 270 valid trials in experiment 2, and 360 valid trials in experiment 3. Praat [13] was used to analyse the voice waveforms for each participant separately in terms of the F_0, sound pressure level, and the F_1 and F_2 under each experimental condition in each experiment. The formants were analysed by the Burg method and the setting of the maximum number of formants was five. 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. We averaged the data for each condition and ran a Tukey–Kramer test (JMP).

3. Results

3.1 Experiment 1

3.1.1 Fundamental frequency

The fundamental frequencies are expressed by converting to cent from Hz, using the formula $1200 \times \log_2 (f1 / f0)$ as the pitch ratio between the actually uttered voice of the fundamental frequency (f1) and the utterance fundamental frequency at which the utterance challenge (f0) (100 cent = 1 semitone). There was no significant difference in the pitch for each condition for each note.

3.1.2 Sound pressure level

Figure 1 shows the mean of the sound pressure levels obtained under the no-noise condition and the four levels of noise in experiment 1. As can be seen in Figure 1, as participants sang while seeing the sound level meter, the sound pressure level increased as the noise increased. However, a significant difference did not appear between each note under each noise level.

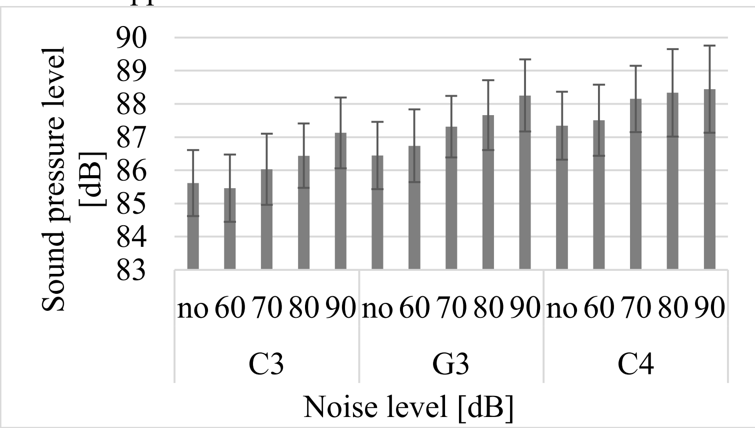


Figure 1: Mean of the sound pressure levels of five stages of noise conditions for each note in experiment 1. Error bars represent the 95 % confidence interval. Each set of five lines from left to right represent the target pitch heights for each noise conditions.

3.1.3 Formant frequency

There was no significant difference in the F_1 and F_2 values for each note under each experimental condition. The F_1 and F_2 values increased as the note became higher. These phenomena could be seen in other experiments (Figures 3 and 6)

3.2 Experiment 2

3.2.1 Fundamental frequency

There was no significant difference in the fundamental frequencies for each condition. The characteristic tendency was not seen either.

3.2.2 Sound pressure level

The sound pressure level decreased as the feedback level was increased for each note (Figure 2). There were significant differences between 85 and 95 dB (p=0.0275), 75 and 95 dB (p=0.0023) in C3, 75 and 85 dB (p=0.0044), 85 and 95 dB (p=0.0003), 75 and 95 dB (p<0.001) in G3, 85 and 95 dB (p<0.001), and 75 and 95 dB (p<0.001) in C4.

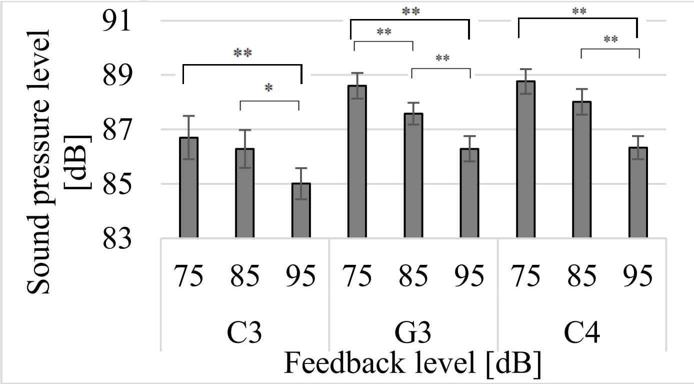


Figure 2: Mean of the sound pressure levels of three stages of feedback conditions for each note in experiment 2. Error bars represent the 95 % confidence interval. Each set of three lines from left to right represent the target pitch heights for each feedback conditions. (*p<0.05, **p<0.01)

3.2.3 Formant frequency

The values of F_1 and F_2 decreased slightly as the feedback level increased for each note (Figure 3). However, there was no significant difference under each experimental condition.

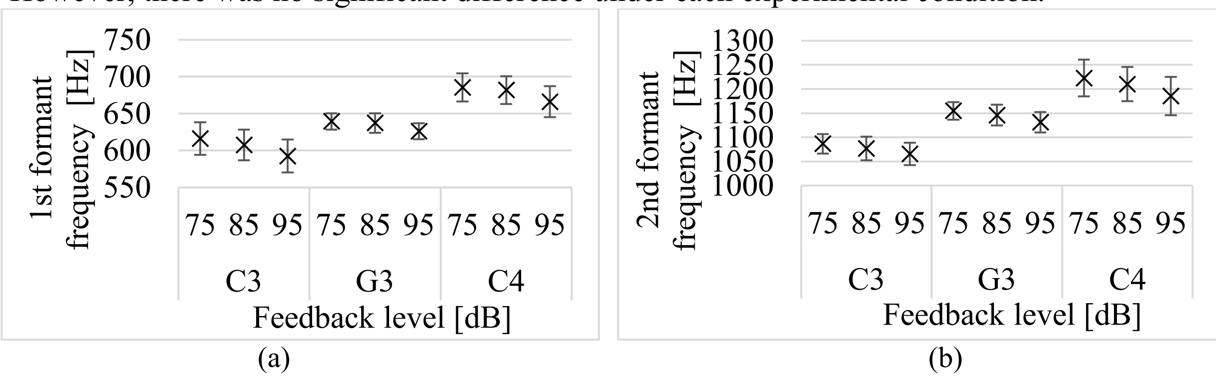


Figure 3: Mean of the F_1 (a) and F_2 (b) of three stages of feedback conditions for each note in experiment 2. Error bars represent the 95 % confidence interval. Each set of three lines from left to right represent the target pitch heights for each feedback conditions.

3.3 Experiment 3

3.3.1 Fundamental frequency

The G3 pitch for which the cut-off frequency was 1,000 Hz was significantly lower than the through condition (p=0.0322) (Figure 4). The pitches when using a 2,000 and 3,000 Hz high-pass filter were lower than the through condition and the pitches obtained using a 1,000 Hz high-pass filter in C3 and C4. However, there was no significant statistical difference found.

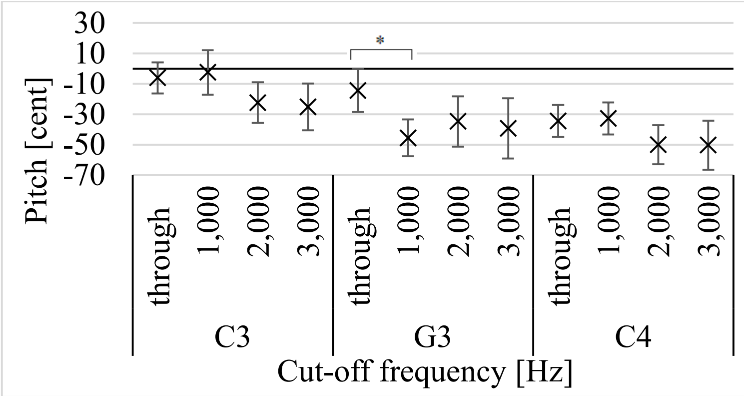


Figure 4: Mean of the fundamental frequencies of the 4 stages of high-pass filtered conditions for each note in experiment 3. Error bars represent the 95 % confidence interval. Thick black line at 0 cent for each set of 4 lines indicates the target pitch heights: C3 (130.81 Hz), G3 (196 Hz), and C4 (261.63 Hz). (*p<0.05)

3.3.2 Sound pressure level

The sound pressure level decreased as the cut-off frequencies increased for each note (Figure 5). There were significant differences between the through and the 3,000 Hz high-pass filtered condition (p=0.0275), the 1,000 Hz and the 3,000 Hz high-pass filtered condition (p=0.0201) in C3, the 1,000 and the 2,000 Hz high-pass filtered condition (p=0.0402), the through and the 2,000 Hz high-pass filtered condition (p=0.0006), the through and 3,000 Hz high-pass filtered condition (p=0.0009) in G3 and the through and the 2,000 Hz high-pass filtered condition (p=0.0002), the through and the 3,000 Hz high-pass filtered condition (p<0.001), the 1,000 Hz and the 2,000 Hz high-pass filtered condition (p=0.0019), and the 1,000 Hz and the 3,000 Hz high-pass filtered condition (p=0.0001).

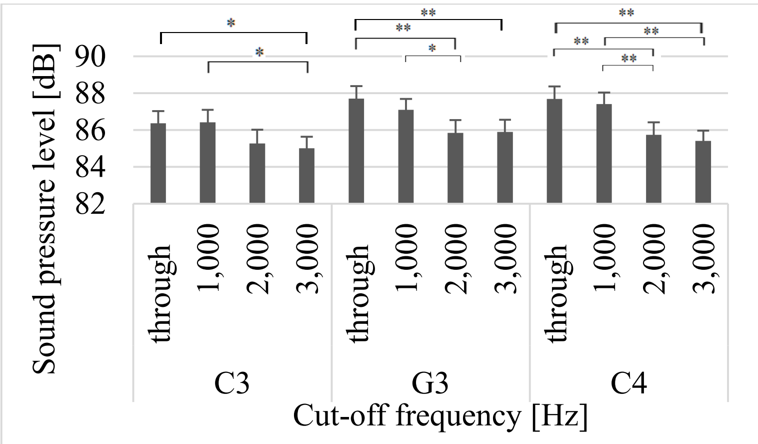
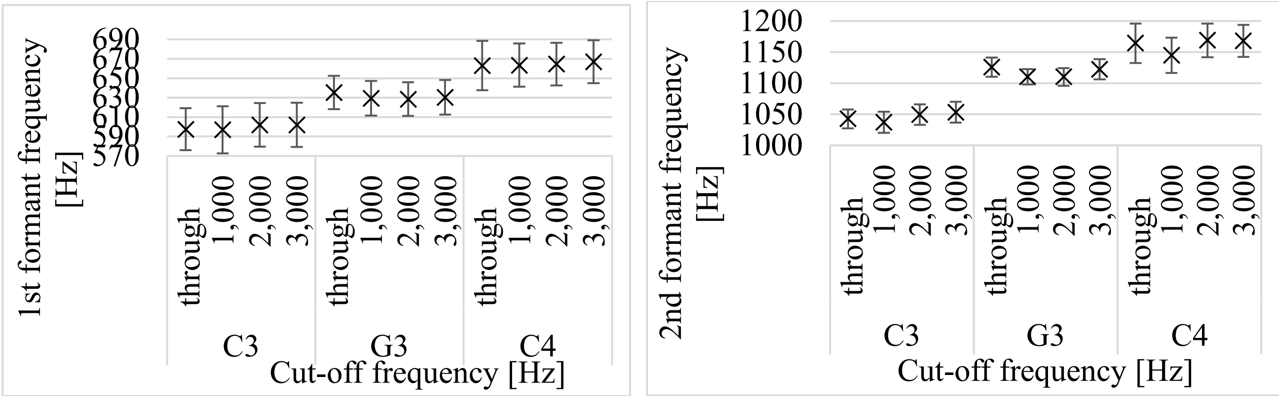


Figure 5: Mean of the sound pressure levels of the 4 stages of high-pass filtered conditions for each note in experiment 3. Error bars represent the 95 % confidence interval. Each set of 4 lines from left to right represent the target pitch heights for each high-pass filtered conditions.

3.3.3 Formant frequency

There was no significant difference in F_1 and F_2 values for each note under the each experimental condition (Figure 6). We obtained almost the same results under the different conditions for each note.



(a) (b)

Figure 6: Mean of the *F1* (a) and *F2* (b) of the 4 stages of high-pass filtered conditions for each note in experiment 3. Error bars represent the 95 % confidence interval. Each set of 4 lines from left to right represent the target pitch heights for each noise conditions.

4. Discussion

We have found more stable pitches in singing than in speech in experiments 1 and 2. In speech production, an increase in F_0 under noise has been widely recognized [1-5]. However, we have not found the same increase of F_0 in experiment 1. In singing, people were required to match their voices with the target pitch heights. It is considered to be the reason why the F_0 in singing is difficult to affect even if people are exposed to noise. Similarly, it was thought that lower voice feedback made F_0 increase. On the contrary, the louder voice feedback made F_0 decrease. However, the result of F_0 in experiment 2 was almost the same as in experiment 1. On the other hand, we have observed the decrease of F_0 when participants received high-pass filtered auditory feedback in experiment 3. In particular, there was a tendency for F_0 to decrease when participants received high-pass filtered auditory feedback with a cut-off frequency over 2,000 Hz. We used the high-pass filter to assess the role of the higher frequencies of a singer's own voice on auditory feedback. Vowel information is not included over 2,000 Hz of vocal sound. The result of experiment 3 indicates that there may be a difference in the effect, whether the vowel information is contained or not. These results suggest that loud noise and quiet or loud voice feedback do not affect the F_0 in singing, but when people receive their own voices emphasized over 2,000 Hz, their F_0 may decrease. With regard to the case of the decreasing F_0, further research may be required.

Correspondingly, it is well known that the vocal sound pressure level increases under noisy environments. In experiment 1, we have observed the tendency the sound pressure level to increase as the noise increases. In previous reports [1-5], the sound pressure level of the speech was free. Therefore, it is thought that the increase of the sound pressure level under noise is a natural occurrence. However, in the present study, as the participants uttered a vowel while watching the sound level meter, the sound pressure level increased as the noise became greater. Therefore, it is considered that the influence of the loudness of the noise is highly significant. Similar results were obtained in experiments 2 and 3. The sound pressure level decreased as the feedback level became greater and as the cut-off frequencies became higher. These results indicate that the feedback level of the participants' own voices has a strong influence on the sound pressure level of their utterances.

Although it is said that the F_1 and the F_2 values increase under noise, the results in experiment 1 have not supported this statement in the pitches at C3 and G3. In C4, the value of F_1 seemed to increase as the noise increased. It is thought further investigation is needed into this facet of singing. In experiment 2, the F_1 and the F_2 values seemed to decrease as the feedback level increased. In experiment 3, the value of F_1 was constant regardless of the increase of the cut-off frequencies. There were differences between the results of experiments 2 and 3 for F_1. However, as it can be seen in Figures 2 and 5, the results of the sound pressure level show a common trend. Therefore, it is considered that the loudness of the feedback voice increases as the cut-off frequency becomes high in experiment 3 as well as when the voice feedback level increases in experiment 2. The difference in these feedback sounds was dictated by whether or not the low frequency of the vocal sound was cut. In experiment 3, when the high-pass filter was used, the participant's F_1 was removed. They did not receive their F_1 loudly; therefore, they maintained their F_1 value. Therefore, it may be possible that the lack of lower frequencies affected the production and the perception of the singing voice. Our hypothesis is that, when people receive loud feedback that is their own voice including F_1, they feel unconsciously that their F_1 is sufficiently loud and then decrease their F_1 automatically. This explains why participants maintained their F_1 in experiment 3.

5. Conclusion

The present study has investigated the effects of Lombard, Fletcher, and the high-pass filtered auditory feedback on singing. The F_0 has not been affected under the Lombard and Fletcher environments. However, for received voice feedback that was emphasized over 2,000 Hz, the value of F_0 has decreased.

As the feedback noise was increased, the trend for the sound pressure level of the participant's voice to increase was observed. When the participants received their own voice loudly, the sound pressure level decreased significantly. It has also been observed that, when the participants received their voice with the higher frequencies emphasized, the sound pressure level decreased significantly.

Finally, the results have indicated the possibility that, if a high enough level of F_1 was perceived by the auditory organs, it would be decreased sensitively.

The results of the present study have indicated that auditory feedback affects the F_0, the sound pressure level and the F_1 of a singing voice.

ACKNOWLEDGMENT

This work was supported by KAWAI FOUNDATION FOR SOUND TECHNOLOGY & MUSIC.

REFERENCES

1 Lombard, E. Le signe de l'elevation de la voix, Annuals Maladies Oreille Larynx, Nez, Pharynx, 37, 101–109, (1911).
2 Korn, T. S. Effect of psychological feedback on conversational noise reduction in rooms, Journal of the Acoustical Society of America, 26, 793–794, (1954).
3 Kano, Y. A study of voice levels in noise for selection of hearing aids, Journal of Otolaryngology of Japan, 88, 138–147, (1985).
4 Uemura, Y. Morise, M., and Nishiura, T. Improvement of speech recognition performance based on the conversion of Lombard features, IEICE Technical Report, SP2010-1, 1–6, (2010).
5 Patel, R., and Schell, K. W. The influence of linguistic content on the Lombard effect, Journal of Speech, Language, and Hearing Research, 51, 209–220, (2008).
6 Fletcher, H. Raff, G. M., and Parmley, F. Study of the effects of different sidetones in the telephone set, Western Electrical Company, Report, (19412), (1918).
7 Lane, H., and Tranel, B. The Lombard sign and the role of hearing in speech, Journal of Speech, Language, and Hearing Research, 14(4), 677–709, (1971).
8 Siegel, G. M., and Pick Jr, H. L. Auditory feedback in the regulation of voice, The Journal of the Acoustical Society of America, 56(5), 1618–1624, (1974).
9 Sundberg, J., The science of the singing voice, Dekalb, IL: Northern Illinois University Press, (1987).
10 Sundberg, J., Lã, F. M., and Gill, B. P. Formant tuning strategies in professional male opera singers, Journal of Voice, 27(3), 278-288, (2013).
11 Sundberg, J. Formant technique in a professional female singer. Acta Acustica united with Acustica, 32(2), 89-96, (1975).
12 Henrich, N., Kiek, M., Smith, J., and Wolfe, J. Resonance strategies used in Bulgarian women's singing style: A pilot study. Logopedics Phoniatrics Vocology, 32(4), 171-177, (2007).