

Article

# Subjective Preferences for Birdsong and Insect Song in Equal Sound Pressure Level

Yoshiharu Soeta \*  and Hiroko KagawaBiomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST),  
Osaka 563-8577, Japan; h-kagawa@aist.go.jp

\* Correspondence: y.soeta@aist.go.jp; Tel.: +81-72-751-8526

Received: 28 December 2019; Accepted: 22 January 2020; Published: 25 January 2020



**Featured Application:** This study is expected to be applied in the improvement of soundscape quality in public spaces, particularly for sound signals that provide visually impaired people with information about basic infrastructure.

**Abstract:** Birdsong is used as a sound signal for visually impaired people in train stations in Japan. However, such sound signals were reportedly difficult to be localized by over 40% of visually impaired people. More than 40% of visually impaired people reported that such sound signals were difficult to identify. The sound pressure level of birdsong is typically low because higher levels of birdsong can cause annoyance to users or residents. Therefore, preferences for birdsong should be investigated. Importantly, birdsong and insect song have been shown to have a positive effect on soundscapes. However, preferences for different types of birdsong and insect song have not been investigated. The current study sought to clarify which types of birdsong and insect song are preferred, and to determine the dominant physical parameters that relate to the preference. We evaluated subjective preferences for various types of birdsong and insect song using paired comparison tests. The relationships between subjective preference and physical parameters were examined. The results indicated that *Horornis diphone* and *Teleogryllus emma* were the most preferred types of birdsong and insect song, both of which are common in Japan. The maximum peak amplitude of the autocorrelation function, determinants of which are pitch salience, loudness, and spectral content, such as centroid, flux, and rolloff, are significant parameters influencing subjective preference.

**Keywords:** soundscape; preference; birdsong; insect song

## 1. Introduction

Many kinds of sound signals are applied to convey information in Japanese public spaces. These signals can be divided into three groups. First, alarm signals are used to provide information about potentially life-threatening events. Second, attention signals are used to provide information about basic infrastructure, such as the locations of stairs, toilets, and ticket gates. Third, announcements are used to convey information via speech.

The guidelines for improving the facilitation of transportation for passenger facilities of public transportation in Japan [1] recommend the use of simulated birdsong for attention signals that inform visually impaired people about the location of stairs in train stations. It is required to adjust the sound pressure levels (SPLs) of the signals to avoid excessive or insufficient volume, and to ensure that they cause no discomfort to users, staff, or surrounding residents. However, a previous study found that more than 40% of visually impaired people reported that such signals were hard to detect [2]. Thus, there is a need for the development of attention signals that are not unpleasant and can be presented at high SPLs.

Introducing birdsong into public spaces has been found to increase the subjective pleasantness of the soundscapes [3]. Among natural sounds, birdsong is reported to be judged as the most effective and beneficial type of sound for improving sound environments [4–6]. Similar to birdsong, insect song has been shown to have a positive effect on soundscapes [5,7]. Thus, birdsong and insect song are less likely to negatively affect sound environments in public spaces, even when presented at high SPLs.

Previous soundscape studies of birdsong and insect song have reported positive effects. However, users' preferences for various types of birdsong and insect song are currently unclear. Specific types of birdsong or insect song may have a more positive impact on sound environments in public spaces. The aim of the current study was to clarify which types of birdsong and/or insect song are preferred, and to determine the dominant physical characteristics related to these preferences.

## 2. Materials and Methods

### 2.1. Subjective Preference Tests

Since the guidelines [1] recommend a duration of 5 s or less for the sound signals, birdsongs and insect songs with one phrase less than 5 s were selected for subjective preference tests. We tested 18 types of birdsong that were recorded using a dummy-head microphone (KU100, Nuemann, Berlin, Germany) at a sampling rate of 48 kHz and a sampling resolution of 24 bits (WILD BIRDS, Victor, Tokyo, Japan). The birdsong stimuli are summarized in Table 1. Figures 1 and 2 show the temporal waveforms and spectrograms of the 18 birdsong stimuli. We tested 16 types of insect song that were recorded using a linear pulse code modulation recorder microphone (LS-10, OLYMPUS, Tokyo, Japan) at a sampling rate of 44.1 kHz and a sampling resolution of 16 bits (Japan NatureSounds Guide, UEDA NATURE SOUND, Tokyo, Japan). The insect song stimuli are summarized in Table 2. Figures 3 and 4 show the temporal waveforms and spectrograms of the 16 insect song stimuli. The duration of the stimuli was between 0.4 and 2.0 s.

**Table 1.** Birdsong stimuli used in this study.

Scientific Name	Order
<i>Halcyon coromanda</i> (HC)	Coraciiformes
Latham (L)	Cuculiformes
<i>Cuculus canorus</i> (CC)	Cuculiformes
<i>Cuculus saturatus</i> (CS)	Cuculiformes
<i>Strix uralensis</i> (SU)	Strigiformes
<i>Otus scops</i> (OS)	Strigiformes
<i>Caprimulgus indicus</i> (CI)	Caprimulgiformes
<i>Streptopelia orientalis</i> (SO)	Columbiformes
<i>Terpsiphone atrocaudata</i> (TA)	Passeriforme
<i>Garrulus glandarius</i> (GG)	Passeriforme
<i>Porzana fusca</i> (PF)	Passeriforme
<i>Parus minor</i> (PM)	Passeriforme
<i>Horornis diphone</i> (HD)	Passeriforme
<i>Zosterops japonicus</i> (ZJ)	Passeriforme
<i>Turdus sibiricus</i> (TS)	Passeriforme
<i>Prunella rubida</i> (PR)	Passeriforme
<i>Eophona personata</i> (EP)	Passeriforme
<i>Emberiza cioides</i> (EC)	Passeriforme

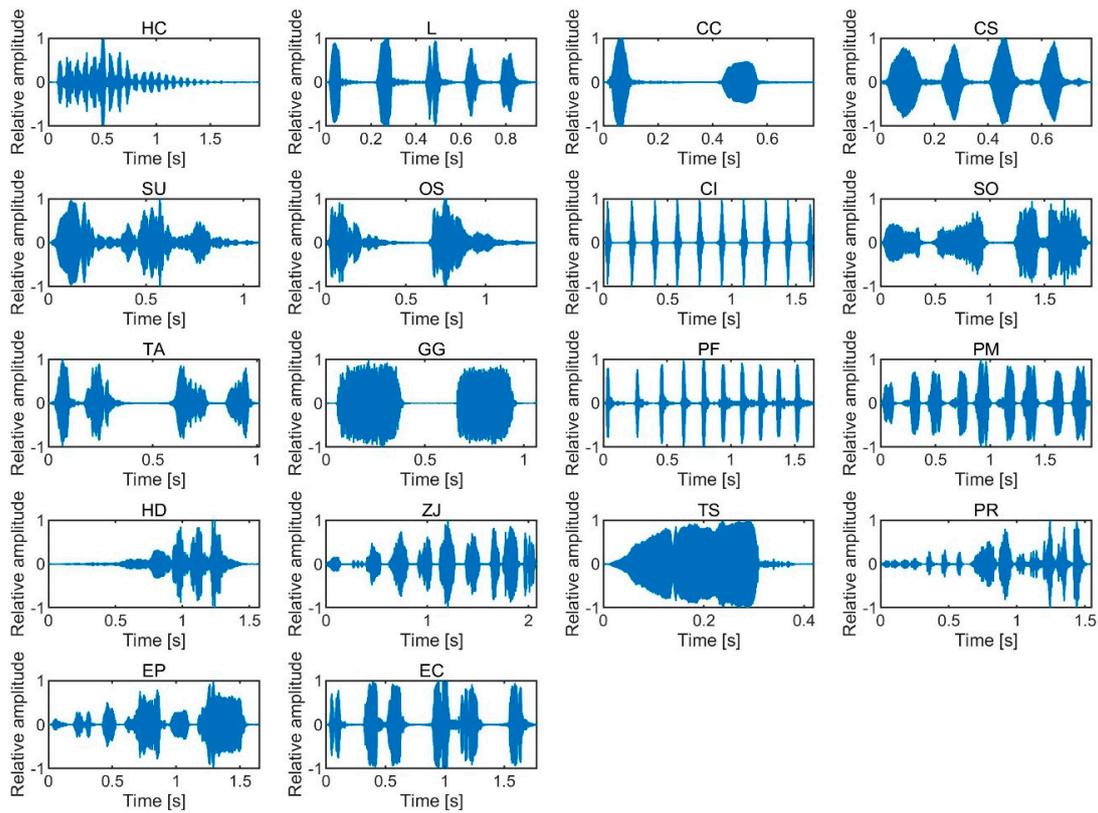


Figure 1. Temporal waveforms of the 18 birdsong stimuli used.

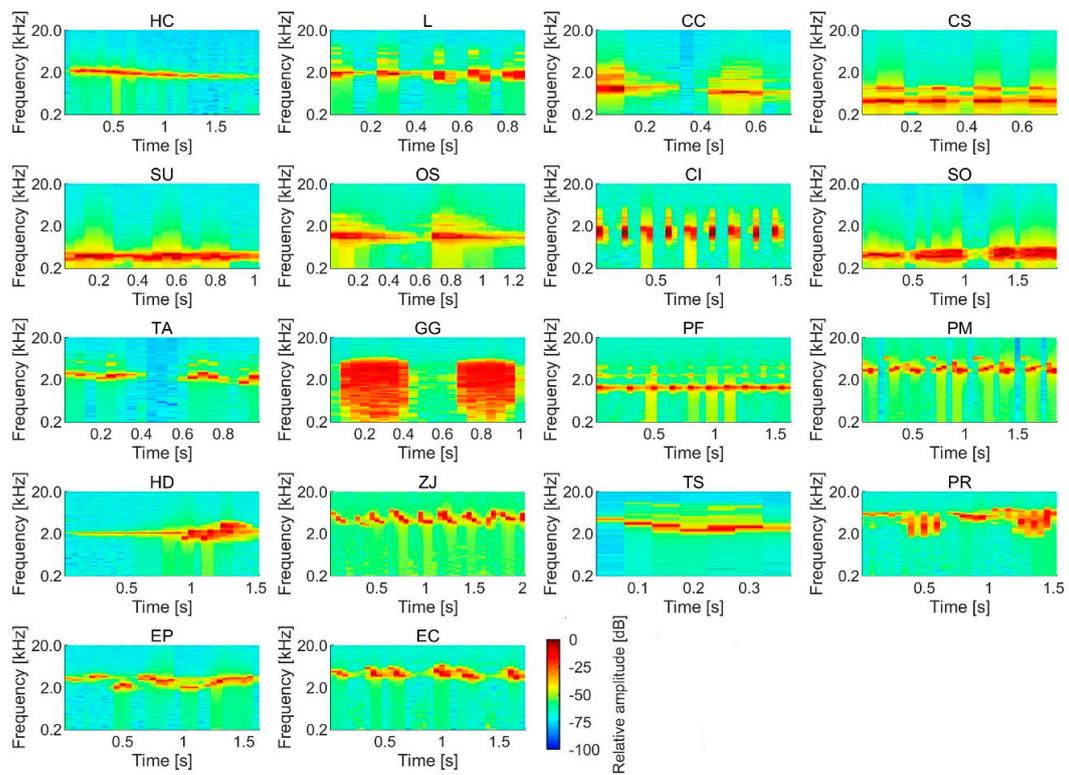


Figure 2. Spectrograms of the 18 birdsong stimuli used.

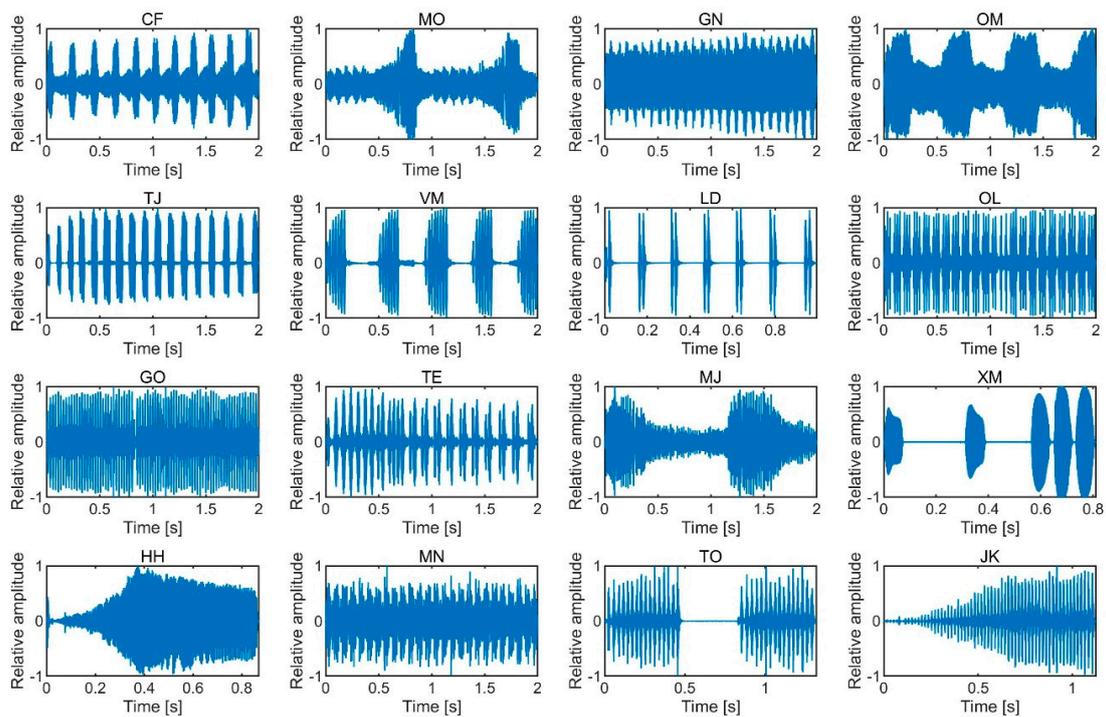


Figure 3. Temporal waveforms of the 16 insect song stimuli used.

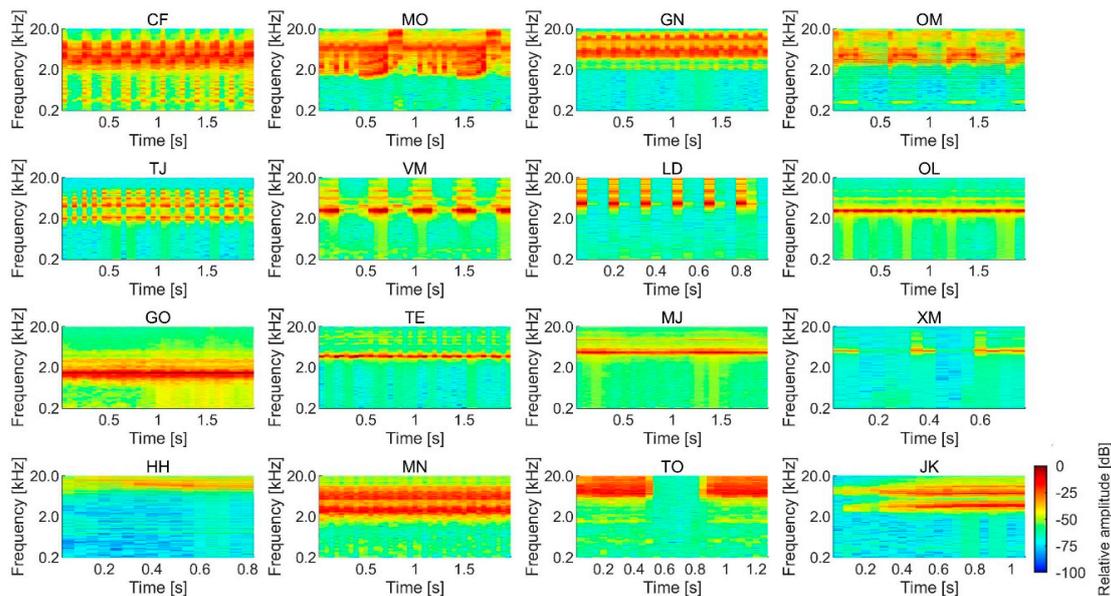


Figure 4. Spectrograms of the 16 insect song stimuli used.

The experiments were conducted in a soundproof and comfortable thermal environment. The stimuli were presented to participants’ ears through a headphone amplifier (HDVD800, Sennheiser, Wedemark, Germany) and headphones (HD800, Sennheiser, Wedemark, Germany). All stimuli were presented at an equivalent continuous A-weighted SPL ( $L_{Aeq}$ ) of 70 dBA, measured over a time period  $T$ , which was the stimulus duration in this experiment. This was because the guidelines [1] recommend that the  $L_{Aeq}$  of the sound signals be at least 10 dB greater than the  $L_{Aeq}$  of the background noise, which is approximately 60 dBA in the train stations of Japan [8].  $L_{Aeq}$  was verified using a dummy head microphone (KU100, Neumann, Berlin, Germany) and a sound calibrator (Type 4231, Brüel & Kjær, Narum, Denmark).

**Table 2.** Insect song stimuli used in this study.

Scientific Name	Order
<i>Cryptotympana facialis</i> (CF)	Hemiptera
<i>Meimuna opalifera</i> (MO)	Hemiptera
<i>Graptosaltria nigrofuscata</i> (GN)	Hemiptera
<i>Oncotympana maculaticollis</i> (OM)	Orthoptera
<i>Tanna japonensis</i> (TJ)	Orthoptera
<i>Velarifictorus micado</i> (VM)	Orthoptera
<i>Loxoblemmus doenitzi</i> (LD)	Orthoptera
<i>Oecanthus longicauda</i> (OL)	Orthoptera
<i>Gryllotalpa orientalis</i> (GO)	Orthoptera
<i>Teleogryllus emma</i> (TE)	Orthoptera
<i>Meloimorpha japonica</i> (MJ)	Orthoptera
<i>Xenogryllus marmoratus</i> (XM)	Orthoptera
<i>Hexacentrus hareyamai</i> (HH)	Orthoptera
<i>Mecopoda niponensis</i> (MN)	Orthoptera
<i>Tettigonia orientalis</i> (TO)	Orthoptera
<i>Japanese katydid</i> (JK)	Orthoptera

Fifteen (eleven males; four females) and sixteen (eleven males; five females) participants with normal hearing and an age range of between 20 and 41 years (median age of 24.5 and 24.4 years and standard deviation of 7.2 and 6.9) took part in the subjective preference for birdsongs and insect songs, respectively. It is considered, from our previous studies [9], that at least ten participants are necessary for the statistical power and generality of the results of these types of psychoacoustic experiments. Informed consent was obtained from each participant after the nature of the study was explained. The study was approved by the ethics committee of the National Institute of Advanced Industrial Science and Technology (AIST) of Japan.

Scheffé's paired comparison methods [10] were conducted for all combinations of pairs (i.e., 153 pairs ( $N[N-1]/2$ ,  $N = 18$ ) for bird songs and 120 pairs ( $N[N-1]/2$ ,  $N = 16$ ) for insect songs) of stimuli by interchanging the order and presenting the pairs in a random order. Scheffé's method assigns one combination to each participant for comparison [10], the modified Scheffé's method assigns all combinations to each participant for comparison, and this is repeated while changing the participant [11,12]. This is appropriate when only a few participants are available. The silent duration between stimuli in each pair was 1.0 s. After the presentation of each pair, the participants were required to judge their preference based on seven-point scorings. The scores of preferences for each participant were defined as the scale values (SVs) of preference. The duration of a single experimental session was approximate 25 to 35 min. Tests for birdsongs and insect songs were conducted separately. The normality assumption of SVs of the preference for each stimulus was verified using the Shapiro–Wilk test [13]. The significance of the main effects was tested by analysis of variance (ANOVA) [10–12].

## 2.2. Physical Parameters

Parameters obtained from the autocorrelation function (ACF) have been proposed for noise and sound quality evaluations [9,14]. To determine ACF parameters, the normalized ACF of the signals recorded by the microphone,  $p(t)$ , is defined by:

$$\varphi(\tau) = \frac{\Phi(\tau)}{\sqrt{\Phi(0)\Phi(0+\tau)}}, \quad (1)$$

where:

$$\Phi(\tau) = \frac{1}{2T} \int_{-T}^{+T} p'(t)p'(t+\tau)dt. \quad (2)$$

$2T$  is the integration interval and  $p'(t) = p(t) * s_e(t)$ , where  $s_e(t)$  is the ear sensitivity corresponding to an A-weighted filtering, for the sake of expedience [9,14]. Normalized ACF is calculated using the geometric mean of the energy at zero and the energy at zero +  $\tau$  to ensure that the normalized ACF satisfies  $0 \leq \varphi(\tau) \leq 1$ .

The ACF parameters are calculated from the normalized ACF.  $\tau_1$  and  $\varphi_1$  are defined as the delay time and the amplitude of the first maximum peak.  $\tau_1$  and  $\varphi_1$  are related to the perceived pitch and the pitch strength of sounds, respectively [9,15]. Higher values of  $\tau_1$  and  $\varphi_1$  indicate that the sound has a lower or a stronger pitch, respectively. The effective duration of the envelope of the normalized ACF,  $\tau_e$ , is defined by the ten-percentile delay and represents a repetitive feature containing the sound source itself [14]. The width of the first decay,  $W_{\varphi(0)}$ , is defined using the delay time interval at a normalized ACF value of 0.5.  $W_{\varphi(0)}$  is equivalent to the spectral centroid [9]. Higher values of  $W_{\varphi(0)}$  indicate that the sound includes a higher proportion of low-frequency components.

Psychoacoustic parameters, such as loudness, sharpness, and roughness, have been applied for sound quality evaluation in previous studies [16]. Loudness is equivalent to the physical strength of a sound, and includes the transfer function of the outer and middle ear and masking effects [17]. Sharpness is a measure of the high-frequency component of a sound. It is calculated using a sound's partial loudness and a weighting function [16]. Roughness quantifies the subjective perception of the fast (15–300 Hz) amplitude variations of a sound. It is calculated based on the excitation pattern in the critical-band filter banks [18].

A range of parameters, such as zero-crossing rate (ZCR), entropy, spectral centroid, and so on, have been introduced for the analysis of audio features [19] and applied to soundscape classification [20, 21]. The ZCR is the number of times the signal changes from positive to negative values or vice versa, divided by the length of signal time. ZCR is considered as a measure of the noisiness of a signal and the spectral characteristics of a signal in a rather coarse manner. The short-term entropy of energy,  $H(i)$ , of the sequence  $e_j$  in the  $K$  sub-frame of fixed duration is defined as:

$$H(i) = - \sum_{j=1}^K e_j \log_2(e_j), \tag{3}$$

where:

$$e_j = \frac{E_{subFramej}}{\sum_{k=1}^K E_{subFramek}}, \tag{4}$$

where  $E_{subframek}$  is the energy for each sub-frame. The entropy of energy is considered as a measure of abrupt changes in a signal. The spectral centroid and spectral spread characterize the spectral position and shape. Spectral centroid,  $C_i$ , is the center of gravity of the spectrum, and the  $C_i$  value of the  $i$ th audio frame is defined as:

$$C_i = \frac{\sum_{k=1}^{W_{fL}} kX_i(k)}{\sum_{k=1}^{W_{fL}} X_i(k)}. \tag{5}$$

Spectral spread,  $S_i$ , is the second central moment of the spectrum and is defined as:

$$S_i = \sqrt{\frac{\sum_{k=1}^{W_{fL}} (k - C_i)^2 X_i(k)}{\sum_{k=1}^{W_{fL}} X_i(k)}}. \tag{6}$$

Spectral entropy measures the peakedness of the spectrum and has been used to discriminate between speech and music. The spectral entropy is calculated in a similar way to the entropy of energy in the frequency domain. Spectral flux,  $F_{(i,i-1)}$ , is a measure of the variability of the spectrum over time and is defined as:

$$FL_{(i,i-1)} = \sum_{k=1}^{W_{fL}} (NE_i(k) - NE_{i-1}(k))^2, \tag{7}$$

where  $NE_i(k)$  is the  $k$ th normalized digital Fourier transform coefficient at the  $i$ th frame. Spectral rolloff is a measure of the skewness of the spectral shape. The spectral rolloff is defined as the frequency below which a certain percentage (usually around 90%) of the distribution is concentrated. The spectral rolloff frequency is usually normalized so that it takes values between 0 and 1.

We calculated  $\tau_1$ ,  $\varphi_1$ ,  $W_{\varphi(0)}$ ,  $\tau_e$ , loudness, sharpness, roughness, ZCR, entropy, spectral centroid, spectral spread, spectral entropy, spectral flux, and spectral rolloff to characterize the birdsongs and insect songs qualitatively. The integration interval or windows length used for the analysis was 0.5 s, and the running step was 0.1 s in all calculations. The average values of physical parameters were used as representative values thereafter. The analyses were carried out using a MATLAB-based program (Mathworks, Natick, MA, USA).

### 2.3. Multiple Regression Analysis

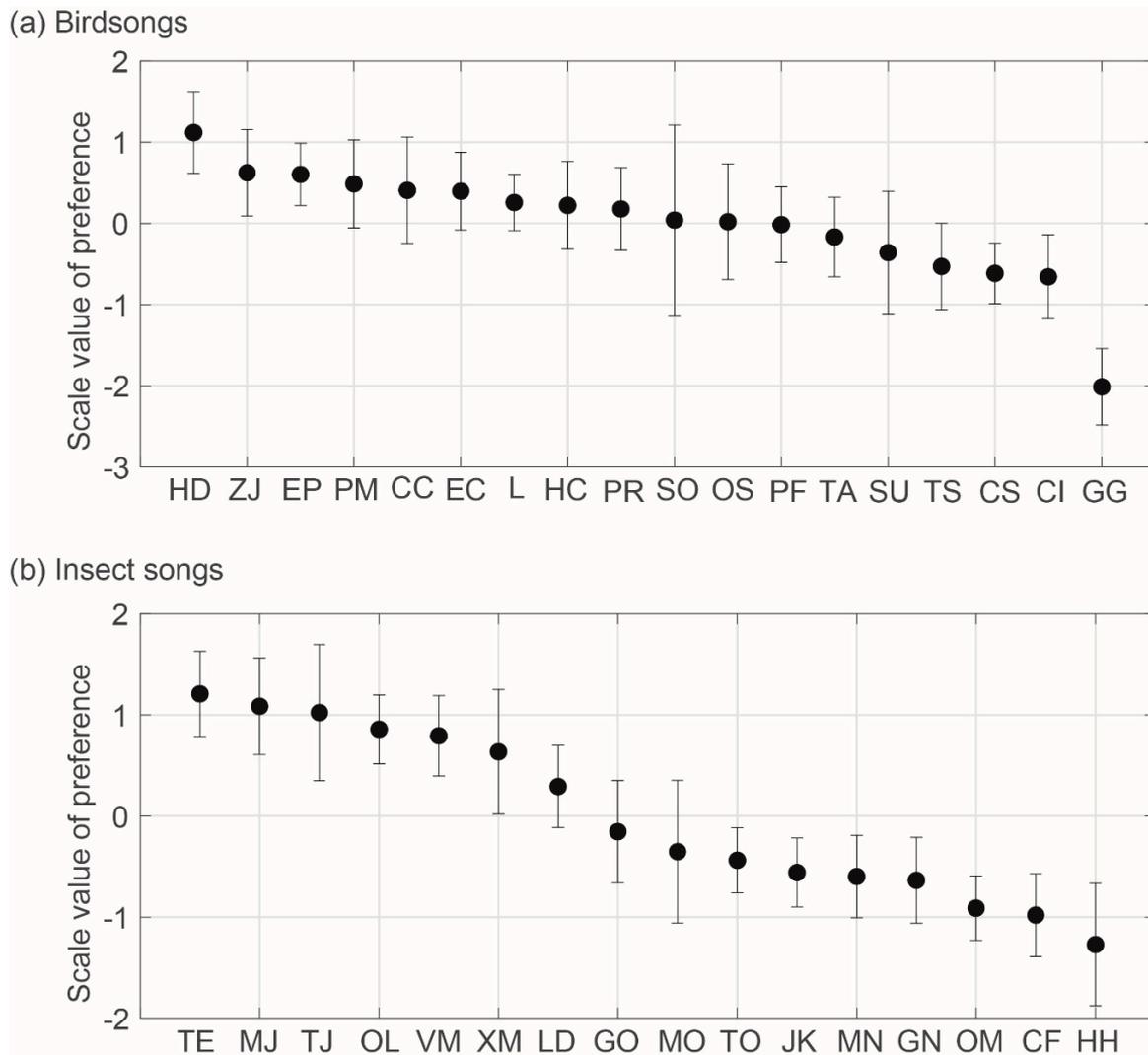
In this stimulus configuration, sharpness and spectral centroid were highly correlated with  $W_{\varphi(0)}$ , ZCR, spectral rolloff, and each other ( $|r| > 0.80$ ,  $p < 0.01$ ). ZCR was highly correlated with spectral entropy and spectral rolloff ( $|r| > 0.70$ ,  $p < 0.01$ ). Therefore, sharpness, spectral centroid, and ZCR were excluded from the multiple regression analysis.

To calculate the effects of each physical parameter on SVs of preference, multiple linear regression analyses were carried out using SPSS statistical analysis software (SPSS version 22.0, IBM Corp., Armonk, NY, USA). The forward-backward stepwise selection was carried out to identify and quantify the significant physical parameters of subjective preference. The criteria applied for entry and removal were based on the significance level of the  $f$ -value and set at 0.05 and 0.10, respectively. Parameters with a variance inflation factor of 10 or higher were excluded to avoid multicollinearity.

## 3. Results and Discussion

ANOVA for the SV of the preference revealed that the main effect (i.e., difference between the stimuli) was statistically significant ( $F(17, 4184) = 295.7$ ,  $p < 0.001$ , for birdsong,  $F(15, 3479) = 395.1$ ,  $p < 0.001$ , for insect song). Figure 5 shows the SV of preferences to birdsong and insect song. The most preferred birdsong was *Horornis diphone*. *Horornis diphone* is one of the three major species of Passeriform in Japan and is considered to produce a beautiful twittering sound. The most preferred insect song was *Teleogryllus emma*. *Teleogryllus emma* is the largest Grylloidea in Japan, known colloquially as a cricket, and its song is often heard throughout Japan in autumn. It is well known that lexical information on words and contextual information on sentences improve speech intelligibility and play an important role in verbal communication [22]. That is, word familiarity plays an important role in intelligibility. Similarly, familiarity with birdsongs and insect songs may affect preferences. Judgments differed among participants regarding preferences for *Streptopelia orientalis*. All participants judged that *Garrulus glandarius* was the least preferred birdsong, producing a sound resembling a creaky voice. Two birdsongs, *Zosterops japonicus* and *Parus minor*, and two insect songs, *Teleogryllus emma* and *Tanna japonensis*, have similar time patterns and scale values of preference, suggesting that time patterns affect preference.

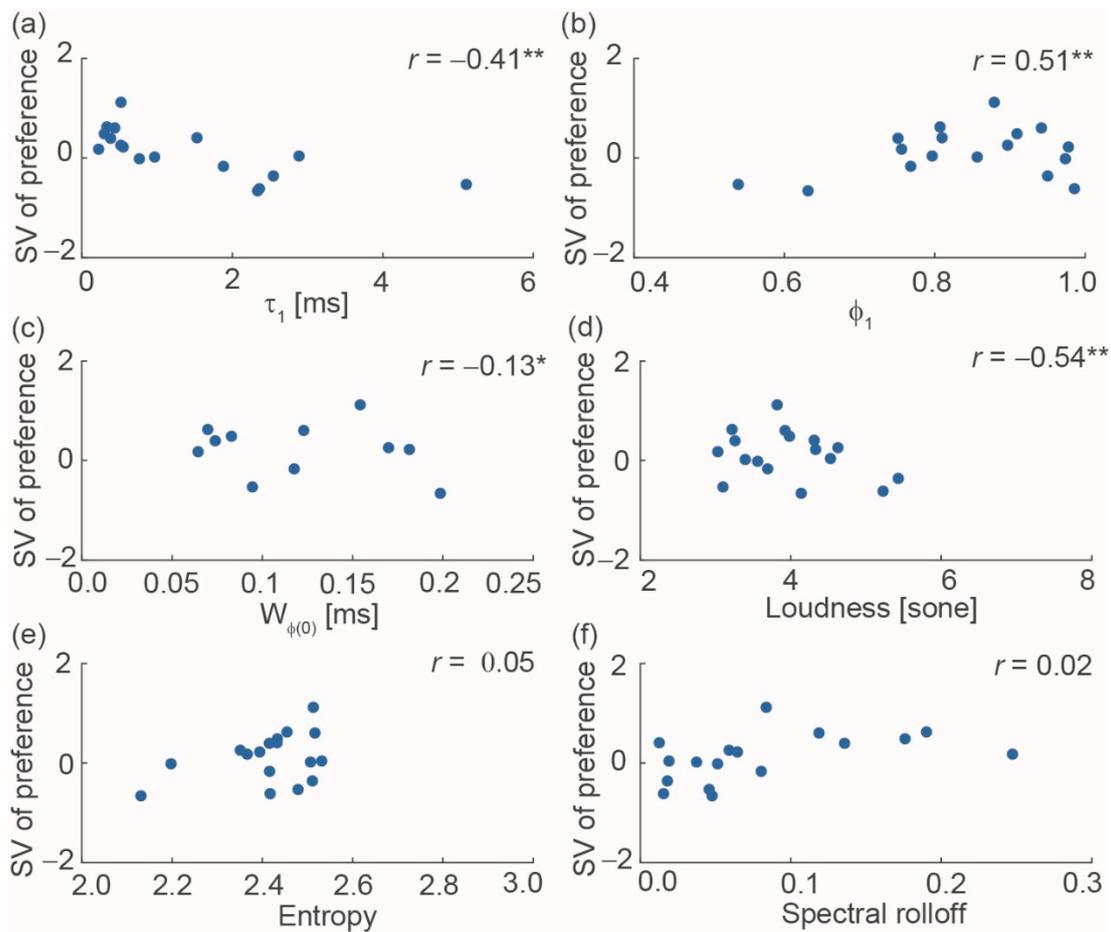
Table 3 shows the correlation coefficients between each physical parameter and subjective preference for birdsongs and insect songs. Figures 6 and 7 show the relationship between representative physical parameter and subjective preference for birdsongs and insect songs. SVs of preference for birdsong and insect song stimuli had a high positive correlation with  $\varphi_1$  ( $r = 0.51$ ,  $p < 0.01$ ;  $r = 0.54$ ,  $p < 0.01$ ), and a high negative correlation with  $\tau_1$  ( $r = -0.41$ ,  $p < 0.01$ ;  $r = -0.37$ ,  $p < 0.01$ ) and loudness ( $r = -0.54$ ,  $p < 0.01$ ;  $r = -0.44$ ,  $p < 0.01$ ). The correlation coefficients between subjective preference and duration of the birdsong and insect song stimuli were  $-0.17$  and  $0.17$ , respectively.



**Figure 5.** Scale values of preferences for (a) birdsong and (b) insect song. Each symbol indicates the mean value and error bars indicate standard deviations.

**Table 3.** Correlation coefficients between each physical parameter and subjective preference for birdsongs and insect songs. Asterisks indicate the level of significance, i.e., \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Parameter	Birdsongs	Insect Songs
$\tau_1$	-0.41 **	-0.37 **
$\varphi_1$	0.51 **	0.54 **
$W_{\varphi(0)}$	-0.13 *	0.25 **
$\tau_e$	-0.28 **	0.02
loudness	-0.54 **	-0.44 **
sharpness	0.21 **	-0.48 **
roughness	0.04	0.14 *
ZCR	0.13 *	-0.64 **
entropy	0.05	-0.49 **
spectral centroid	0.20 **	-0.54 **
spectral spread	0.02	0.37 **
spectral entropy	-0.04	-0.59 **
spectral rolloff	0.02	0.47 **
spectral flux	0.14 *	-0.68 **

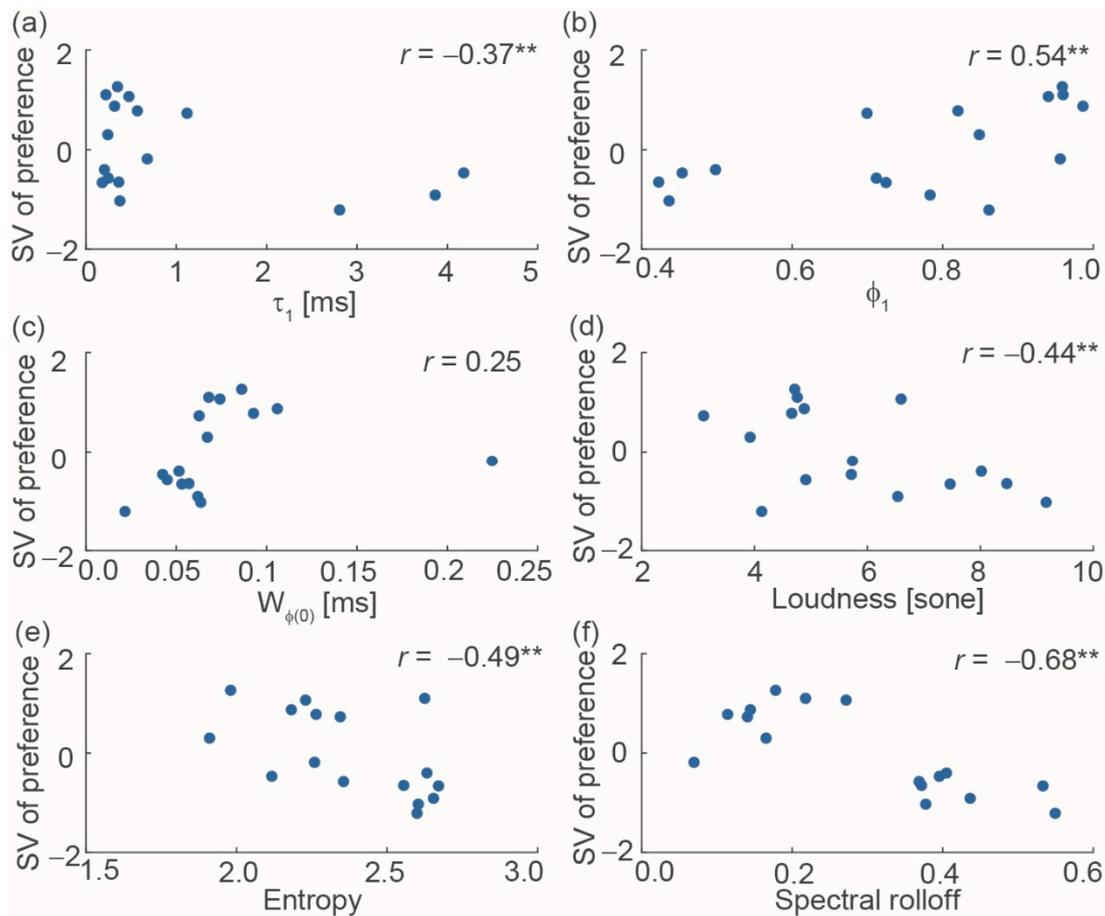


**Figure 6.** Relationships between physical parameters and scale values of preference for birdsong stimuli. (a)  $\tau_1$ , (b)  $\phi_1$ , (c)  $W_{\phi(0)}$ , (d) loudness, (e) entropy, and (f) spectral rolloff. The numbers represent correlation coefficient values. Asterisks represent the level of significance, i.e.,  $** p < 0.01$ ,  $* p < 0.05$ .

Sounds with higher  $\phi_1$  have a salient pitch [15] and noises with lower  $\phi_1$ , such as those generated by a refrigerator [23] or an air conditioner [24], tend to cause more annoyance. These findings suggest that birdsong and insect song stimuli with a salient pitch, which have  $\phi_1$  values around 0.9, are preferable. It has been reported that the tonal component, which is correlated with  $\phi_1$  value, causes annoyance [25]. Although this is not consistent with the present findings, it can be explained by periodic variations in the amplitude and frequency of birdsongs and insect songs. The vibrato-like variation can decrease annoyance.

Since lower  $\tau_1$  indicates higher pitch, a negative correlation between subjective preference and  $\tau_1$  values suggests that birdsongs and insect songs with a higher pitch are preferable. However, this finding is not consistent with previous findings regarding subjective annoyance caused by refrigerator noises [23]. These findings suggest that sounds with a higher and more salient pitch are preferred, whereas sounds with a higher but unnoticeable pitch cause annoyance.

The negative correlations between subjective preference and loudness meant that quieter birdsong stimuli were preferable. Since SPL was fixed at 70 dBA in the current experiment, loudness may be a useful index for identifying birdsong stimuli that are pleasant when presented at high SPLs.



**Figure 7.** Relationships between physical parameters and scale values of preference for birdsong stimuli. (a)  $\tau_1$ , (b)  $\phi_1$ , (c)  $W_{\phi(0)}$ , (d) loudness, (e) entropy, and (f) spectral rolloff. The numbers represent correlation coefficient values. Asterisks represent the level of significance, i.e., \*\*  $p < 0.01$ , \*  $p < 0.05$ .

Insect song stimuli with lower entropy were preferred, suggesting that abrupt changes in SPL were not preferred. Lower spectral entropy and rolloff were preferred, suggesting that insect song stimuli with broader bandwidth were preferred. Thus, spectral content appeared to play a more important role in determining preferences for insect song stimuli than for birdsong stimuli.

A multiple linear regression analysis was performed with the SVs of preference for birdsong and insect song stimuli as the outcome variable. The final model showed that  $\phi_1$ ,  $\tau_1$ , loudness, and entropy were significant parameters for birdsong stimuli, while  $\phi_1$ ,  $W_{\phi(0)}$ ,  $\tau_e$ , loudness, and spectral rolloff were significant parameters for insect song stimuli:

$$SV_{\text{birdsong}} \approx b_0 + b_1 * \phi_1 + b_2 * \tau_1 + b_3 * \text{loudness} + b_4 * \text{entropy}, \quad (8)$$

$$SV_{\text{insect song}} \approx i_0 + i_1 * \phi_1 + i_2 * W_{\phi(0)} + i_3 * \tau_e + i_4 * \text{loudness} + i_5 * \text{spectral rolloff}. \quad (9)$$

The ANOVA indicated the statistical significance of the model ( $F(4, 265) = 63.33, p < 0.001$ , for birdsong,  $F(5, 250) = 101.84, p < 0.001$ , for insect song). The adjusted coefficient of determination,  $R^2$ , was 0.54 for birdsong and 0.66 for insect song. The standardized partial regression coefficients in Equations (8) and (9) are summarized in Table 4.

**Table 4.** Significant predictive parameters and standardized partial regression coefficients in multiple linear regression analyses for birdsong and insect song. Asterisks indicate the level of significance, i.e., \*\*  $p < 0.01$ , \*  $p < 0.05$ .

	Predictive Parameter	Standardized Partial Regression Coefficients
Birdsong	$b_1, \varphi_1$ **	0.27
	$b_2, \tau_1$ **	-0.23
	$b_3, \text{loudness}$ **	-0.45
	$b_4, \text{entropy}$ **	0.21
Insect song	$i_1, \varphi_1$ **	0.54
	$i_2, W_{\varphi(0)}$ **	-0.63
	$i_3, \tau_e$ *	-0.09
	$i_4, \text{loudness}$ **	0.33
	$i_5, \text{spectral rolloff}$ **	-1.01

The maximum peak amplitude of the ACF,  $\varphi_1$ , which signifies pitch salience and loudness were significant predictive variables for both birdsong and insect song stimuli. Although the partial regression coefficients of loudness for birdsong stimuli were negative, those for insect song stimuli were positive. The simple correlation between subjective preference and loudness was negative. Therefore, this difference suggests that the contributions from predictive variables other than loudness were much larger for insect song stimuli than for birdsong stimuli. Similar to the simple correlation analysis, spectral rolloff was also found to be a significant factor for the prediction of preferences for insect song stimuli.

The width of the first decay of the ACF,  $W_{\varphi(0)}$ , which is equivalent to the spectral centroid, exhibited a larger and negative standardized partial regression coefficient for insect song stimuli. However,  $W_{\varphi(0)}$ , was not significant for birdsong stimuli. This finding could be explained by pitch strength. Birdsong has relatively higher  $\varphi_1$  values, meaning that birdsong stimuli have a stronger pitch; in this case, pitch is mainly judged by the delay time of the maximum peak of the ACF,  $\tau_1$ . Insect song stimuli have relatively lower  $\varphi_1$  values, meaning that insect song stimuli have a weaker pitch; in this case, pitch is mainly judged by  $W_{\varphi(0)}$ . Similar results were previously obtained in a study of pitch sensation for airplane noise [26].

#### 4. Conclusions

In the current study, preferable birdsong and insect song stimuli and the dominant physical parameters related to preferences were investigated. The results indicated that bird and insect songs that are commonly found in Japan, *Horornis diphone* and *Teleogryllus emma*, were rated as the most preferable bird and insect song stimuli. Preferences for birdsong and insect song stimuli may differ in other countries. The maximum peak amplitude of the ACF,  $\varphi_1$ , and loudness were significant predictive variables for both birdsong and insect song stimuli. Pitch variables related to the parameters of delay time of the maximum peak of the ACF,  $\tau_1$  for birdsong and the width of the first decay of the ACF,  $W_{\varphi(0)}$ , for insect song, were significant parameters. Spectral rolloff, which measured the skewness of the spectral shape, provided the largest contribution to subjective preference for insect songs. These findings gave insights into the future design of attention signals to utilize physical parameters, such as  $\varphi_1$ ,  $\tau_1$  and loudness. For example, attention signals with higher  $\varphi_1$  and lower  $\tau_1$  and loudness values are preferable in equal  $L_{Aeq}$  conditions.

The current findings suggest that it may be useful for future research to develop attention signals that are preferable, even when presented at high SPLs. Current security guidelines in Japan recommend the use of simulated birdsong stimuli as attention signals rather than real birdsong stimuli. Based on the current results, combining the advantages of birdsong A and insect song B may enable the development of a more pleasing sound. In addition, the present findings need validation by participants who are visually impaired, because there were no visually impaired people among the participants.

**Author Contributions:** Conceptualization, Y.S.; methodology, Y.S.; formal analysis, H.K.; investigation, Y.S.; resources, Y.S.; data curation, Y.S. and H.K.; writing—original draft preparation, Y.S.; writing—review and editing, Y.S.; visualization, Y.S.; project administration, Y.S.; funding acquisition, Y.S. All authors have read and agree to the published version of the manuscript.

**Funding:** This work was partly supported by a Grant-in-Aid for Scientific Research (B) (Grant No. 18H03324) from the Japan Society for the Promotion of Science.

**Acknowledgments:** We thank Benjamin Knight, MSc, from Edanz Group (<https://en-author-services.edanzgroup.com/>) for editing a draft of this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ministry of Land, Infrastructure, Transport and Tourism. *Guidelines for Improving the Facilitation of Transportation for Passenger Facilities of Public Transportation*; Ministry of Land, Infrastructure, Transport and Tourism: Tokyo, Japan, 2007. (In Japanese)
2. Foundation for Promoting Personal Mobility and Ecological Transportation. *Basic Case Study Report on Movement Support for Traffic Hub*; Foundation for Promoting Personal Mobility and Ecological Transportation: Tokyo, Japan, 2009. (In Japanese)
3. De Coensel, B.; Vanwetswinkel, S.; Botteldooren, D. Effects of natural sounds on the perception of road traffic noise. *J. Acoust. Soc. Am.* **2011**, *129*, EL148–EL153. [[CrossRef](#)] [[PubMed](#)]
4. Guastavino, C. The ideal urban soundscape: Investigating the sound quality of French cities. *Acta Acust. Acust.* **2006**, *92*, 945–951.
5. Jeon, J.Y.; Lee, P.J.; You, J.; Kang, J. Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds. *J. Acoust. Soc. Am.* **2010**, *127*, 1357–1366. [[CrossRef](#)] [[PubMed](#)]
6. Hong, J.Y.; Jeon, J.Y. Designing sound and visual components for enhancement of urban soundscapes. *J. Acoust. Soc. Am.* **2013**, *134*, 2026–2036. [[CrossRef](#)] [[PubMed](#)]
7. Ou, D.; Mak, C.M.; Pan, S. A method for assessing soundscape in urban parks based on the service quality measurement models. *Appl. Acoust.* **2017**, *127*, 184–193. [[CrossRef](#)]
8. Shimokura, R.; Soeta, Y. Characteristics of train noise in ground and underground stations with side and island platforms. *J. Sound Vib.* **2011**, *330*, 1621–1633. [[CrossRef](#)]
9. Soeta, Y.; Ando, Y. *Neurally Based Measurement and Evaluation of Environmental Noise*; Springer: Tokyo, Japan, 2015.
10. Scheffé, H. An analysis of variance for paired comparisons. *J. Am. Statist. Assoc.* **1952**, *147*, 381–400.
11. Sato, S. *Statistical Method of Sensory Test*; JUSE Press: Tokyo, Japan, 1985. (In Japanese)
12. Nagasawa, S. Improvement of the Scheffé's method for paired comparisons. *Kansei Eng. J.* **2002**, *3*, 47–56. [[CrossRef](#)]
13. Shapiro, S.S.; Wilk, M.B. An analysis of variance test for normality (complete samples). *Biometrika* **1965**, *52*, 591–611. [[CrossRef](#)]
14. Ando, Y.; Cariani, P. *Auditory and Visual Sensations*; Springer: New York, NY, USA, 2009.
15. Yost, W.A. Pitch strength of iterated ripple noise. *J. Acoust. Soc. Am.* **1996**, *100*, 3329. [[CrossRef](#)] [[PubMed](#)]
16. Zwicker, E.; Fastl, H. *Psychoacoustics: Facts and Models*; Springer-Verlag: Berlin, Germany, 1999.
17. Zwicker, E. Procedure for calculating loudness of temporally variable sounds. *J. Acoust. Soc. Am.* **1977**, *62*, 675–682. [[CrossRef](#)] [[PubMed](#)]
18. Daniel, P.; Weber, R. Psychoacoustical roughness: Implementation of an optimized model. *Acta Acust. United Ac.* **1997**, *83*, 113–123.
19. Giannakopoulos, T.; Pikrakis, A. *Introduction to Audio Analysis: A MATLAB® Approach*; Academic Press: Oxford, UK, 2008.
20. Alías, F.; Socoró, J.C.; Valero, X. A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds. *Appl. Sci.* **2016**, *6*, 143. [[CrossRef](#)]
21. Giannakopoulos, T.; Spyrou, E.; Perantonis, S. Recognition of Urban Sound Events Using Deep Context-Aware Feature Extractors and Handcrafted Features. In *Artificial Intelligence Applications and Innovations*, 2nd ed.; MacIntyre, J., Maglogiannis, I., Iliadis, L., Pimenidis, E., Eds.; Springer: Cham, Switzerland, 2019; pp. 184–195.

22. Amano, S. Effects of lexicon and coarticulation on phoneme perception. *J. Acoust. Soc. Jpn. (E)* **1993**, *14*, 91–97. [[CrossRef](#)]
23. Sato, S.; You, J.; Jeon, J.Y. Sound quality characteristics of refrigerator noise in real living environments with relation to psychoacoustical and autocorrelation function parameters. *J. Acoust. Soc. Am.* **2007**, *122*, 314–325. [[CrossRef](#)] [[PubMed](#)]
24. Soeta, Y.; Shimokura, R. Sound quality evaluation of air-conditioner noise based on factors of the autocorrelation function. *Appl. Acoust.* **2017**, *124*, 11–19. [[CrossRef](#)]
25. Hansen, H.; Verhey, J.L.; Weber, Reinhard. The magnitude of tonal content. A review. *Acta Acust. United Ac.* **2011**, *97*, 355–363. [[CrossRef](#)]
26. Soeta, Y.; Kagawa, H. Three dimensional psychological evaluation of aircraft noise and prediction by physical parameters. *Build. Environ.* **2020**, *16*, 106445. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).