human disturbance in general raises animal anxiety in the presence of humans, which in turn leads to difficulties in monitoring wild individuals. Since at least certain vocalisations are species specific (Hartwig 2005, Rupell 2009) and individual specific (Bouchet et al. 2012, Neumann et al. 2010, Geissman and Nijman 2006), knowing a species repertoire and an individual’s call pattern will help population monitoring activities for conservation purposes. It has been demonstrated by Adi et al. (2009) that a population census in Emberiza hortulana using vocal recognition can predict the population density and structure simpler and more accurate as compared to other population monitoring methods.

Furthermore, different types of vocalisations might represent the internal emotion of animals either positive state (Boissy et al. 2007) or negative state (Dupjan et al. 2008, Fichtel et al. 2001), so the understanding of macaque's vocalization will be useful for the species’ management and to ensure their welfare.

The aim of this study therefore is to provide quantitative data of crested macaque vocalizations to get a better understanding about the communication system of this species. The study also attempt to categorize calls based on the contexts in which they were emitted and examine whether acoustic features overlap within and between calls in different contexts to provide an estimation of the graded nature of the crested macaque vocal communication system. In addition to support species management and conservation, this research will provide a key of future studies on the vocal communication of Sulawesi macaques.

2 LITERATURE REVIEW

Mechanism of Vocal Production of Primates

Vocal production of mammals employs several organs such as the lungs, larynx, and vocal tract (Fitch and Hauser 1995). Lungs are the source of power; they convert the steady air into a series of puff air. The larynx then modulates air pressure to become a sound source through the opening and closing process of the larynx valve. The larynx valve is under the control of the vocal folds (vocal cords). There are two major mechanisms of vocal sound production: phonation and impulsive or noisy sound. Phonation occurs when the vocal folds vibration through larynx produces repetitive series of puff air in a certain level of frequency, which is the fundamental frequency (Figure 2). Impulsive (cough and clicks) or noisy sounds (hisses and whispers) are busted out by the high pressure of air that forces the opening of the larynx. The variation of primate vocalizations results from the phonation of the sound source. The sound source then flows through the supralaryngeal vocal tract on its way out to the lips and nostrils. This organ includes the pharyngeal, oral, and nasal cavities, which serve to resonate the sound source. The differences in the length and shape of the human vocal tract are known to cause the difference of resonance frequency (known as formants in human, Fitch and Hauser 1995). Non-human primates have a longer and thinner tongue, larynges that are positioned higher in the neck, and there is a relative lack of flexible soft tissues in the supralaryngeal cavities compared to humans (Owren
and Linker 1995). These differences have presumably limited the ability of vocal variation in non-human primate.

The air puff from the larynx produces sound, which can be either tonal or noisy sound. Tonal sound is characterized by one excitation frequency, which is sometimes followed by harmonics or overtones. The harmonics are integer multiple of the fundamental frequency. While noisy sound is a periodic signal, which may also be produced at the vocal folds as a result of air turbulence. Noisy call can be produce either at the larynx or above the larynx (due to constriction elsewhere in the vocal tract, Fitch and Hauser 1995). The combination of tonal and non-tonal call can be called complex calls (Micheletta 2012).

**Sound Transmission**

The sound waves are only able to propagate through a medium. The sound waves that travel through gasses and liquid are called longitudinal waves while sound waves that transmit through solid medium are called transversal waves (Bradbury and Vehrencamp 1998). Hence the vocal exchange in primate is a longitudinal sound wave that is transmitted through air. The sound wave transmission is simply described as the following explanation; when there is no vibration from any sound source, the air pressure is called ambient pressure. The vibration of the sound source increases the collision of molecules in the air. When the molecules move and stay away due to the collision, the air pressure is below ambient pressure and increase again when the next tremor occurs. The state when the molecule assembled is called compression while the state when the molecule expanded is called rarefaction. The temporal pattern of rise and fall in signal amplitude is called the waveform of the signal. The largest distance of air pressure toward ambient pressure is called amplitude. The distance between two successive peaks of high pressure called wavelength. The number of cycles (one rise and fall of the signal) per second is called frequency. The unit of frequency, cycles/second is called Hertz. The time is needed by a cycle is called period (Bradbury and Vehrencamp 1998).

**The Constraint of Sound Propagation in the Field**

Acoustic transmission can be constrained by two fundamental problems: attenuation and degradation, which affect signal quality. Attenuation can be caused by the atmospheric absorption, ground attenuation, scattering of a sound beam, and deflection of sound by stratified media (Wiley and Richards 1978). Signal attenuation in the dense forest is relatively severer than in open habitat. It also occurs over long distance and low transmission height (Maciej et al. 2011). Sound degradation results from irregular amplitude fluctuation (e.g. atmospheric turbulence of the wind) and reverberation (e.g. vegetation). In an open habitat, irregular amplitude fluctuation happens often and masks the low frequency of amplitude modulation. In the forest habitat, reverberation is more severe than in open habitat and masks the transmission of high frequencies. Both types of degradation affect more to the temporal pattern of amplitude or intensity modulation than frequency modulation (Wiley and Richards 1978).
The constraints of sound wave transmission in the field can affect the quality of sound recordings (Maciej et al. 2011). Signal quality can also be masked by background noise such as the sound of other animals or observer. Large distances between sound emitter and observer will potentially enhance sound degradation. Some simple efforts such as pointing the microphone towards the sound source within close proximity, can overcome the problem. Depending on the call, e.g. low amplitude calls like grunts, good sound quality may be obtained from 3-5 m distance while high amplitude signals might be recorded at 15 m (Fischer et al. 2013). Nevertheless, for further analysis using computers, preliminary assessment of the reliability of recordings needs to be done for any signals recorded in large distance (Maciej et al. 2011).

**Sound Analysis**

Fischer et al. (2013) provided a detailed explanation of the study of bioacoustic in primates along with a guideline for playback experiments. Data collected in the field is transferred into a computer for analysis. Computer program analyses that are widely used to extract acoustic features of sound signal are Avisoft SAS Lab (R. Specht, Berlin), RAVEN (Cornell Lab of Ornithology), PRAAT (Institute of Phonetic Science, http://www.praat.org) or Signal (Engineering Design, Belmont, MA). The computer-aided analysis still faces problems related to the complexity of signal, high variation of vocalization, environmental effect and distance and spatial orientation of the animal towards the microphone. Using a multi-parameter approach, as provided by the software LMA (Schrader and Hammerschmidt 1997), can reduce the importance of these problems.

Sound is a complex signal, mostly made up of non-pure sinusoidal waveforms. To deal with this acoustic complexity, the Fourier transformation technique is required. The main principle of this technique is to transform any continuous waveform into a pure sine wave. Sound signals in pure sinusoidal form are easy to compare or to analyse their frequency, amplitude and relative phase. In a computer program, the Fast Fourier Transform (FFT) breaks down the signal into smaller unit samples and records the amplitude at each point. The advantage of this technique is that sound signals can be analysed rapidly. The disadvantage is that when the number of signal points is too small, the computer will have difficulties to determine the frequency of the signal. This problem can be solved by applying the Nyquist-Shanon equation in which the sampling rate of recording equipment is set at twice the frequency of the signal frequency (Bradburry and Vehrencamp 1998). The sampling rate for most mammal vocal recordings can be set at 40 kHz, which is twice the maximum capability of human hearing, 20 kHz. Yet, the sampling rate that is commonly used is 44.1 kHz, parallel to the sampling rate of a signal in compact disc storage.

The acoustic signal can also be plotted in spectrograms. A spectrogram is a graphical illustration of the distribution of signal energy along a time axis. In a two-dimensional spectrogram, the vertical axis refers to frequency value, while the horizontal axis refers to time (Bradburry and Vehrencamp 1998, Figure 2), while the shading represents amplitude. Spectrograms with good resolution are modified by the choice of frequency range, FFT-size, length of time segment, and